



Universidad Nacional
SAN LUIS GONZAGA



[Reconocimiento-CompartirIgual 4.0 Internacional](https://creativecommons.org/licenses/by-sa/4.0/)

Esta licencia permite a otras combinar, retocar, y crear a partir de su obra, incluso con fines comerciales, siempre y cuando den crédito y licencia a las nuevas creaciones bajo los mismos términos. Esta licencia suele ser comparada con las licencias copyleft de software libre y de código abierto. Todas las nuevas obras basadas en la suya portarán la misma licencia, así que cualesquiera obras derivadas permitirán también uso comercial.

<http://creativecommons.org/licenses/by-sa/4.0/>



UNIVERSIDAD NACIONAL "SAN LUIS GONZAGA"



ESCUELA DE POSGRADO

EVALUACION DE ORIGINALIDAD

CONSTANCIA

El que suscribe, deja constancia que se ha realizado el análisis con el software de verificación de similitud al **BORRADOR DE TESIS** cuyo título es:

"MODELO DE PREDICCIÓN CLIMATOLÓGICA CON INTELIGENCIA ARTIFICIAL (AI) EN LA REGIÓN ICA, 2022"

Presentado por:

QUISPE GALINDO LUCY

Del **DOCTORADO EN GESTIÓN AMBIENTAL**.

Que, se ha recibido del operador del programa informático evaluador de originalidad de la Escuela de Posgrado de la UNICA, el informe automatizado de originalidad, el mismo que concluye de la siguiente manera:

El documento de investigación APRUEBA los criterios de originalidad con un porcentaje de similitud de 3%.

Para dar fe, se adjunta al presente el reporte de similitud de las bases de datos de iThenticate. En Ica 22 de marzo de 2024

Atentamente


UNIVERSIDAD NACIONAL "SAN LUIS GONZAGA"
ESCUELA DE POSGRADO
Dr. LUIS ALBERTO PECHO TATAJE
Director (e)

UNIVERSIDAD NACIONAL "SAN LUIS GONZAGA"
VICERRECTORADO DE INVESTIGACIÓN
ESCUELA DE POSGRADO
DOCTORADO EN GESTION AMBIENTAL



TESIS

**MODELO DE PREDICCIÓN CLIMATOLÓGICA CON INTELIGENCIA
ARTIFICIAL (AI) EN LA REGIÓN ICA, 2022**

LÍNEA DE INVESTIGACIÓN:

RECURSOS HIDRICOS, RIESGOS DE DESASTRES Y CAMBIO CLIMATICO

PRESENTADO POR:

Mag. Lucy Quispe Galindo

GRADO A OBTENER: DOCTOR

ASESOR:

Dr. PEDRO CORDOVA MENDOZA

ICA – PERÚ

2024

DEDICATORIA

El presente trabajo de grado va dedicado a **Dios**, quien como guía estuvo presente en el caminar de mi vida y dándome fuerzas para continuar con mis metas trazadas.

A mi **madre Serafina**, por ser el pilar más importante y por demostrarme siempre el cariño y apoyo incondicional corrigiendo mis faltas y celebrando mis triunfos.

A mi **padre Tomas**, por haberme apoyado en todo momento, por sus consejos, sus valores por la motivación constante.

A mi **hija Lucia**, mi motivación y mi alegría

A mis **hermanos, Lidia y Héctor**, porque me han brindado su apoyo incondicional y por compartir conmigo buenos y malos momentos.

A mi **esposo Amancio**, por apoyarme para continuar y que, con su amor y respaldo, me ayuda alcanzar mis objetivos

A **mis compañeros de aula**, por ser un equipo muy especial y divertido

AGRADECIMIENTO

Quiero expresar mi **gratitud a Dios**, quien con su bendición llena siempre mi vida y a toda mi familia por estar siempre presentes.

De igual manera mis **agradecimientos A las autoridades de la Universidad Nacional San Luis Gonzaga de Ica**, a la escuela de pos grado, a mis maestros en especial, quienes con la enseñanza de sus valiosos conocimientos hicieron que pueda crecer día a día como profesional, gracias a cada una de ustedes por su paciencia, dedicación, apoyo incondicional y amistad.

Finalmente quiero expresar mi más grande y sincero agradecimiento al **Dr. Pedro Córdova, asesor de tesis**, principal colaborador durante todo este proceso, quien con su dirección, conocimiento, enseñanza y colaboración permitió el desarrollo de este trabajo.

ÍNDICE DE CONTENIDOS

INFORME FINAL DE INVESTIGACION	ii
DEDICATORIA	iii
AGRADECIMIENTO	iv
ÍNDICE DE CONTENIDOS	v
ÍNDICE DE TABLAS	vii
ÍNDICE DE FIGURAS	viii
RESUMEN	x
ABSTRACT	xi
I. INTRODUCCIÓN	1
1.1. Aspectos científicos de la investigación	1
1.2. Antecedentes.....	4
1.2.1. Antecedentes internacionales	4
1.2.2. Antecedentes nacionales	5
1.3. Bases Teóricas	6
1.3.1. Teoría modelo ARIMA.....	6
1.3.2. Modelo ARIMA(p, d, q).....	7
1.3.3. Pasos para el Análisis ARIMA.....	7
1.3.4. Estación meteorológica.....	8
1.3.5. Información meteorológica	9
1.3.6. Inteligencia artificial.....	10
1.3.7. Aprendizaje Automático	10
1.3.8. Clima	12
1.4. Realidad problemática	12
1.4.1. Problema general.....	14
1.4.2. Problemas específicos.....	14
1.5. Objetivos de la investigación	14
1.5.1. Objetivo general	14
1.5.2. Objetivos Específicos	14
1.6. Hipótesis de la investigación.....	15
1.6.1. Hipótesis general.....	15
1.6.2. Hipótesis Específicas	15
1.6.3. Variables.....	15
1.7. Justificación e importancia de la investigación	16
1.7.1. Justificación	16
1.7.2. Importancia	17

II. ESTRATEGIA METODOLÓGICA	19
2.1. Área de estudio	19
2.2. Tipo, nivel y diseño de investigación	20
2.2.1. Tipo de investigación.....	21
2.2.2. Nivel de investigación	21
2.2.3. Diseño de la investigación	21
2.3. Población y muestra	21
2.3.1. Población	21
2.3.2. Muestra.....	21
2.4. Técnicas de recolección de datos	22
2.5. Instrumentos de recolección de datos	22
2.6. Técnicas de procesamiento de datos.....	22
2.7. Análisis e interpretación de datos.....	23
III. RESULTADOS	24
3.1. Establecer el modelamiento de los parámetros meteorológicos y la inteligencia artificial en el clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022	24
3.2. Analizar el elemento temperatura y la inteligencia artificial en el clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022. 31	31
3.3. Predecir los parámetros meteorológicos y la inteligencia artificial en los fenómenos atmosféricos en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022.....	3
IV. DISCUSIÓN.....	10
4.1. Discusión de resultados: modelamiento de los parámetros meteorológicos y la inteligencia artificial en el clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022.....	10
4.2. Discusión de resultados: elemento temperatura y la inteligencia artificial en el clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022. 11	11
4.3. Discusión de resultados: los parámetros meteorológicos y la inteligencia artificial en los fenómenos atmosféricos en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022.....	11
V. CONCLUSIONES	13
VI. RECOMENDACIONES	15
VII. REFERENCIAS BIBLIOGRÁFICAS	17
VIII. ANEXOS.....	¡Error! Marcador no definido.

ÍNDICE DE TABLAS

TABLA 1 “CLASIFICACIÓN DE LAS ESTACIONES METEOROLÓGICAS”[21]	9
TABLA 2 RESULTADOS DE LA EVALUACIÓN DE LOS MODELOS DE PREDICCIÓN	9

ÍNDICE DE FIGURAS

Figura 1 Inteligencia artificial y sus ramas [2].....	2
Figura 2 “Aquatinter Single Layer Feedforward neural network (SLFN)”[28]	11
Figura 3 Ciudad de Ica. Departamento de Ica	19
Figura 4 Área donde se encuentra la estación meteorológica, ubicada en el campus de la Universidad Nacional “San Luis Gonzaga” (Av. IOs maestros S/N).....	20
Figura 5 Modelo inicial de comprensión de los datos.....	25
Figura 6 Carga del archivo clima.csv	26
Figura 7 Selección del archivo clima.csv.....	26
Figura 8 Asignación del tipo de dato a las columnas.....	27
Figura 9 Tabla con los datos y sus tipos de datos asignados	27
Figura 10 Information del dataSet Clima	28
Figura 11 Datos del clima	28
Figura 12 Modelo para comprensión gráfica	29
Figura 13 Comprensión gráfica de los datos	30
Figura 14 Transformación de las variables Temperatura y Radiación.....	31
Figura 15 Transformación de la variable Temperatura.....	32
Figura 16 Datos transformados de la Temperatura.....	33
Figura 17 Ampliación del modelo con las variables velocidad y dirección del viento	33
Figura 18 Conversión de la velocidad del viento	34
Figura 19 Resultados del cálculo.....	35
Figura 20 Selección de la variable obtenida en el cálculo	35
Figura 21 Modelo final de transformación de las variables meteorológicas	36
Figura 22 Obtención de la dataSet Clima_H.csv	1
Figura 23 datos del dataSet Clima_H.csv	1
Figura 24 Modelo para análisis gráfico del archivo Clima_H.csv.....	2
Figura 25 Gráfica del comportamiento de la variable de temperatura	2
Figura 26 Comprobación de datos faltantes	4
Figura 27 Verificación de los datos faltantes	5
Figura 28 Comprobación de datos faltantes del primer rango de fechas	5
Figura 29 Selección de registro de datos a imputar	6
Figura 30 Datos seleccionados para el primer periodo	7
Figura 31 Modelo final de inputación de datos	1
Figura 32 Datos concatenados.....	1

Figura 33 Visualización gráfica de puntos para la temperatura.....	2
Figura 34 Comportamiento de las variables meteorológicas.....	2
Figura 35 Modelo de evaluación para los algoritmos de series de tiempo.....	4
Figura 36 Selección de variables de predicción	4
Figura 37 Configuración serie de tiempo.....	4
Figura 38 Configuración ARIMA Model	5
Figura 39 Configuración VAR Model.....	5
Figura 40 Resultados de la evaluación de los modelos ARIMA Model y VAR Model.....	6
Figura 41 Predicción del modelo ARIMA Model	6
Figura 42 Predicción de las temperaturas	7
Figura 43 Predicción de la Radiación	7
Figura 44 Predicción de la velocidad del viento	8
Figura 45 Predicción de la dirección del viento	8

RESUMEN

Objetivo, Este estudio se propuso “establecer el modelamiento de los parámetros meteorológicos que mediante inteligencia artificial contribuye significativamente en el clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022”. **Estrategia metodológica,** adoptada siguió la estructura CRISP-DM y comprendió la preparación de datos, análisis de la temperatura, aplicación de modelos ARIMA y VAR para predicción meteorológica, y evaluación de la eficacia de los modelos. **Resultados,** destacaron la presencia de datos faltantes, cuya imputación fue esencial para mantener la integridad temporal del conjunto de datos. La aplicación de modelos ARIMA y VAR mostró que ARIMA superó en precisión a VAR en varias métricas de evaluación. **Discusión,** se centró en la importancia de abordar los datos faltantes y la necesidad de explorar modelos más avanzados. **Conclusión,** resalta la aplicabilidad de la inteligencia artificial en la predicción climática y sugiere la implementación de sistemas de alerta temprana, con el propósito de enriquecer la gestión de riesgos climáticos en la región de Ica.

Palabras Claves: *Meteorología; Inteligencia Artificial; ARIMA Model; Datos Faltantes; Predicción Climática*

ABSTRACT

Objective, this study aimed to establish the modeling of meteorological parameters that, through artificial intelligence, contribute significantly to the climate at the Research Center for the Study of Solar Activity and its Effects on the Earth, Ica, 2019-2022. **Methodological strategy** adopted followed the CRISP-DM structure and included data preparation, temperature analysis, application of ARIMA and VAR models for weather prediction, and evaluation of the effectiveness of the models. **Results**, highlighted the presence of missing data, whose imputation was essential to maintain the temporal integrity of the data set. The application of ARIMA and VAR models showed that ARIMA outperformed VAR in accuracy in several evaluation metrics. **Discussion**, focused on the importance of addressing missing data and the need to explore more advanced models. **Conclusion**, highlights the applicability of artificial intelligence in climate prediction and suggests the implementation of early warning systems, with the aim of improving climate risk management in the Ica region.

Keywords: Meteorology; Artificial intelligence; ARIMAModel; Missing data; Climate Prediction

I. INTRODUCCIÓN

1.1. Aspectos científicos de la investigación

Estamos llevando a cabo un proyecto en el que empleamos inteligencia artificial y redes neuronales para predecir las condiciones meteorológicas en la región de Ica. Esto se debe a la necesidad de comprender mejor el cambio climático en nuestra área. En nuestra región, disponemos de múltiples estaciones meteorológicas ubicadas en lugares estratégicos, como San Camilo, Hacienda Bernales, Pampa Blanca, Rio Grande, Ocucaje, Huamani, Fonagro (Chincha), San Juan de Yanac y Copara, que proporcionan datos esenciales para este propósito.

El SENAMHI ha estado recopilando datos en estas estaciones sobre las temperaturas máximas y mínimas, así como sobre las precipitaciones. Es importante destacar que en nuestra región, las precipitaciones son bastante escasas.

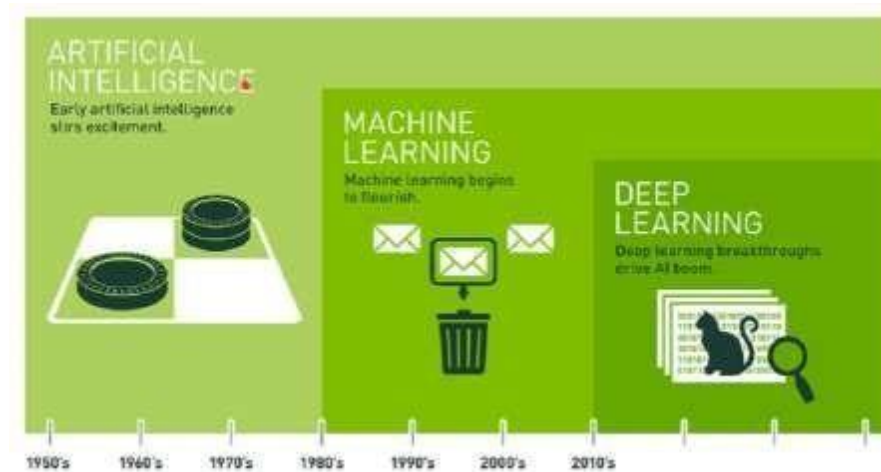
Las variables meteorológicas que afectan al clima de una región además de la temperatura se incluyen otras como la velocidad del viento, la presión etc.

La gran cantidad de información que se viene generando no está siendo aprovechada por las instituciones para tomar decisiones según corresponda a su actividad como el agro, y el propio sistema meteorológico del SENAMHI.

En la actualidad la inteligencia artificial está siendo muy utilizada en el tratamiento de datos a con la finalidad de realizar predicciones para nuestro caso sobre las variables que inciden en el clima de la región y de ser posible tomar decisiones en provecho de toda la ciudadanía.

En su libro '2084 Inteligencia Artificial y el Futuro de la Humanidad', Jhon McCarthy (1956, citado en *Lennox Jhon*, 2021) proporciona una definición de la inteligencia artificial. La describe como la disciplina científica y técnica centrada en la creación de máquinas con capacidades inteligentes. En la actualidad, este término se aplica tanto a

las máquinas inteligentes en sí como a la ciencia y la tecnología que trabajan para desarrollar estas máquinas[1].



Fuente: Udemy: Machine Learning y Data Science con Python

Figura 1 Inteligencia artificial y sus ramas[2]

Con el propósito de adentrarnos en el ámbito de nuestra investigación, la Figura 1, según **Garrido-Peña**, ilustra que la Inteligencia Artificial se centra principalmente en el Machine Learning o aprendizaje automático. Este enfoque se define según Arthur Samuel (1959, citado en Ng, 2020) La disciplina que permite a las computadoras adquirir la habilidad de aprender. sin necesidad de programación explícita'. Una versión más contemporánea de esta definición proviene de Tom Mitchell (1998, citado en Ng, 2020), quien expresa que un programa informático es reconocido por aprender de la experiencia E. en relación a una tarea T y una medida de rendimiento P se relacionan con el desempeño en la tarea T, evaluado mediante P, mejora con la experiencia E[2].

Según **Andrew Ng**, quien imparte el curso 'Machine Learning' en Coursera de la Universidad de Stanford, el campo “del Machine Learning abarca el subcampo del Deep Learning o aprendizaje profundo”. Lo distintivo de este enfoque es la utilización de redes neuronales para procesar datos con algoritmos de mayor precisión. Las redes neuronales, diseñadas para imitar el funcionamiento del cerebro humano, tuvieron un gran auge en las décadas de los 80 y 90, pero su popularidad decreció hacia finales de los 90 debido al alto costo de procesamiento requerido. Sin embargo, experimentaron un resurgimiento en la última década gracias al aumento en la capacidad de cómputo

de las computadoras, lo que permitió llevar a cabo procesos altamente complejos con algoritmos y modelos sumamente eficientes[3].

El clima está constituido por una serie de elementos y factores atmosféricos, y desempeña un papel crucial en la vida cotidiana y en una amplia gama de actividades, como la industria, la agricultura, la ganadería y muchas otras. Su análisis y estudio son esenciales para comprenderlo y para facilitar las diversas actividades cotidianas[31].

Castro et al. señalan que los datos meteorológicos ofrecen numerosas posibilidades en el ámbito de las Geociencias, especialmente en la mejora de la predicción de condiciones meteorológicas futuras. El modelado de datos meteorológicos utilizando redes neuronales profundas se ha transformado en un ámbito de investigación de gran relevancia[4].

Según *Hatfield et al.*, los emuladores de redes neuronales cumplen una función esencial en la parametrización física de los modelos numéricos utilizados en la predicción meteorológica[5]. La información meteorológica se adquiere a través de datos de series temporales, lo que ha llevado a la exploración de modelos de última generación, como los LSTM y TCN. Estos modelos son enfoques especializados de redes neuronales utilizados en la predicción meteorológica, como explican detalladamente los investigadores *Hewage et al.*,[6].

Según la investigación de *Kim et al.*, se está empleando la aplicación de redes neuronales artificiales (RNA) para prever las condiciones de las carreteras en días lluviosos. Este enfoque tiene como objetivo ofrecer a los conductores información valiosa que pueda contribuir a evitar accidentes de tráfico debidos a las malas condiciones de la carretera.[7]. La utilización de *redes neuronales* LSTM (Long Short-Term Memory) mejora la planificación en el sector de generación al prever la insolación solar y la velocidad del viento. Se emplea un modelo clásico de red neuronal con memoria a corto plazo para anticipar la carga eléctrica y la insolación solar combinadas durante el día en hogares con sistemas de energía fotovoltaica en el tejado. Esto se basa en la clasificación meteorológica[8].

Ren et al. destacan la importancia de la predicción meteorológica en la anticipación temprana de los efectos del clima en varios aspectos de la vida humana[9]. *Song y Roh* explican la creación de un emulador de red neuronal (NN) diseñado para parametrizar la radiación. Este emulador se integrará en un modelo de previsión meteorológica operativa utilizado por la Administración Meteorológica de Corea[10].

1.2. Antecedentes

1.2.1. Antecedentes internacionales

Según *Nsabagwa et al.*, la observación de las condiciones meteorológicas durante un período prolongado es lo que caracteriza el clima de una región. Continuando con su explicación, señalan que el clima representa un promedio de los diversos elementos meteorológicos observados a lo largo de un extenso período de tiempo, con la Organización Meteorológica Mundial recomendando un período de referencia de 30 años para esta evaluación[11]. *Dube et al.*, a lo largo del último siglo, se ha evidenciado un incremento global en las temperaturas, así como episodios de lluvias intensas, lo que ha desencadenado un aumento significativo en la variación de inundaciones y sequías en diversas regiones del mundo. Este fenómeno se ha convertido en una preocupación de relevancia mundial en la actualidad[12]. *Mubiru et al.*, en ciertas áreas de África Oriental, las precipitaciones estacionales tienden a llegar de manera anticipada o retrasada, se distribuyen de manera irregular y con frecuencia se registran por debajo de los niveles normales[13].

Lwasa et al., adaptación a cambios en los patrones climáticos, se han propuesto numerosos enfoques de adaptación. Uno de estos enfoques implica mejorar la precisión de las predicciones meteorológicas[14]. *Nsabagwa et al.*, lograr una mayor precisión en los datos meteorológicos requiere la implementación de instrumentos meteorológicos resistentes y una mayor densidad de redes de estaciones meteorológicas. Esto se traduce en lecturas más precisas y una representación mejorada de las áreas bajo observación [11].

Kreuzer et al., dada la considerable demanda, existen diversos modelos disponibles para generar pronósticos de temperatura. En el ámbito de la predicción meteorológica y climática, se encuentran instituciones como los institutos meteorológicos nacionales, entre los que se incluyen el Deutscher Wetterdienst en Alemania y el Meteorológico Noruego, por mencionar algunos ejemplos[15].

Stegmaier et al., no obstante, incluso en estas investigaciones, los usuarios a veces se perciben como un elemento externo en lugar de ser considerados como una componente esencial del proceso de desarrollo de servicios climáticos[15]. **Dumas et al.**, “la falta de un nivel adecuado de conocimiento climático por parte de los usuarios, como su familiaridad con tipos de datos, tecnologías, productos y partes interesadas, a menudo los excluye del proceso de desarrollo de los servicios climáticos en sí”[16]. **Cavelier et al.**, esta situación puede dar lugar a una paradoja en lo que respecta al deseo de servicios climáticos orientados por los usuarios. También respalda la idea de adoptar un enfoque basado en el mercado para involucrar a una diversidad de participantes en el desarrollo de servicios climáticos[16].

1.2.2. Antecedentes nacionales

Vilca, señala que los sistemas neuro-difusos, también conocidos como las redes neuronales y los sistemas difusos son dos técnicas de aprendizaje automático que tienen fortalezas y debilidades diferentes. Las redes neuronales son buenas para aprender de datos complejos, pero pueden ser difíciles de interpretar. Los sistemas difusos son más fáciles de interpretar, pero pueden ser menos precisos.[17].

Según **Armas**, en algunas compañías peruanas, se están observando procesos de transformación tecnológica, digitalización y automatización, que involucran el uso de la Inteligencia Artificial. La investigación se centra en explorar y describir, desde un enfoque cualitativo, cómo se está empleando la Inteligencia Artificial en Perú y las redes neuronales y los sistemas difusos

tienen el potencial de crear nuevos puestos de trabajo y eliminar otros. En general, se espera que estas tecnologías tengan un impacto positivo en el empleo, creando puestos de trabajo en áreas como la ingeniería, la ciencia de datos y la inteligencia artificial. Sin embargo, también es posible que se pierdan puestos de trabajo en áreas que puedan ser automatizadas por estas tecnologías[18].

Contribuye *Uscumayta*, que la “Inteligencia Artificial” (IA) ha sido considerada como una fuerza disruptiva desde hace un tiempo considerable. De hecho, hay quienes anticipan que, en un futuro cercano, la IA será tan fundamental como la electricidad, es decir, que gran parte de nuestra dependencia estará centrada en ella. A pesar de estas perspectivas, la realidad actual de la IA es todavía bastante modesta y se limita en gran medida a brindar soporte y asistencia en diversas actividades humanas[19].

1.3. Bases Teóricas

1.3.1. Teoría modelo ARIMA

El modelo ARIMA constituye una técnica estadística que integra la información proveniente de la autorregresión (AR) y el promedio móvil (MA) con el propósito de modelar series temporales que son estacionarias o estacionalmente estacionarias. Además, ARIMA incorpora el componente de integración (I) para abordar tendencias no estacionarias presentes en los datos[20].

- **Autorregresión (AR):** Corresponde a la relación lineal entre la observación actual y múltiples observaciones anteriores con un cierto retraso. Un modelo AR(p) toma en cuenta las últimas p observaciones.
- **Promedio Móvil (MA):** Refleja la relación lineal entre la observación actual y los errores residuales pasados en un modelo de retraso. Un modelo MA(q) considera los últimos q errores residuales.

- **Integración (I):** Involucra la diferenciación de la serie de tiempo para convertirla en estacionaria. Se considera que una serie de tiempo es estacionaria cuando sus propiedades estadísticas permanecen constantes a lo largo del tiempo[20].

1.3.2. Modelo ARIMA(p, d, q)

Un modelo ARIMA (p, d, q) es un modelo de series temporales que se utiliza para pronosticar valores futuros de una variable a partir de sus valores históricos. El modelo está formado por tres componentes:

p: Orden de la componente de autorregresión (AR).

d: Orden de la componente de integración (I).

q: Orden de la componente de promedio móvil (MA).

1.3.3. Pasos para el Análisis ARIMA

El análisis ARIMA (Autorregresivo Integrado de Medias Móviles) implica varios pasos, que se describen a continuación:

Preprocesamiento de datos:

- ✓ Recopilar datos de series temporales relevantes.
- ✓ Verificar la estacionariedad de los datos.
- ✓ Si los datos no son estacionarios, aplicar diferenciación para hacerlos estacionarios.

Identificación del orden (p, d, q):

p (Orden de Autorregresión): Identificar el número de retrasos necesarios para explicar la estructura de dependencia en la serie temporal.

d (Orden de Integración): Determinar la cantidad de diferenciación requerida para hacer que la serie temporal sea estacionaria.

q (Orden de Medias Móviles): Identificar el número de términos de medias móviles necesarios para explicar la estructura de dependencia en la serie temporal.

Ajuste del modelo ARIMA:

Utilizar los parámetros obtenidos (p , d , q) para ajustar el modelo ARIMA a los datos.

Diagnóstico del modelo:

- ✓ Verificar la validez del modelo a través de pruebas de residuos.
- ✓ Asegurarse de que los residuos sean ruido blanco (sin patrones discernibles).

Predicciones y evaluación:

- ✓ Realizar predicciones utilizando el modelo ajustado.
- ✓ Evaluar la precisión de las predicciones utilizando métricas como el “Error Cuadrático Medio (RMSE) o el Error Absoluto Medio (MAE)”.

Visualización:

- ✓ Representar gráficamente las predicciones en comparación con los datos observados.
- ✓ Evaluar visualmente la bondad del ajuste del modelo.

Estos pasos proporcionan una guía general para realizar un análisis ARIMA. Cabe destacar que la elección adecuada de los parámetros (p , d , q) y la interpretación de los resultados requieren experiencia y conocimientos en series temporales. Además, en la práctica, se pueden realizar ajustes iterativos para mejorar la calidad del modelo [20].

1.3.4. Estación meteorológica

Castillo, en su investigación, menciona que la Organización Meteorológica Mundial (OMM) clasifica las estaciones meteorológicas en función de su finalidad y características específicas. [21].

TABLA 1 “CLASIFICACIÓN DE LAS ESTACIONES METEOROLÓGICAS”[21]

Según su finalidad	Clasificación
“Sinóptica	“Climatológica Agrícolas Especiales Aeronáuticas Satélites
De acuerdo a la magnitud de las observaciones	Principales Ordinarias Auxiliares o adicionales
Por el nivel de observación	Superficie Altitud
Según el lugar de observación” [21]	Terrestre Aéreas Marítimas”[21]

Peña en su investigación sobre “las observaciones realizadas con métodos y en forma sistemática, uniforme, ininterrumpida y a horas establecidas, permiten conocer las características y variaciones de los elementos atmosféricos, los cuales constituyen los datos básicos que utilizan los servicios meteorológicos, tanto en tiempo real como diferido”[22].

Según *Zapata*, sobre la estación meteorológica, “¿Qué es una estación meteorológica?, considera que: Es el lugar donde se realizan mediciones y observaciones puntuales de los diferentes parámetros meteorológicos utilizando los instrumentos adecuados para así poder establecer el comportamiento atmosférico”[23].

1.3.5. Información meteorológica

Campetella et al., “la información meteorológica y los productos, posteriormente, desarrollados, como el pronóstico del tiempo, permiten planificar actividades futuras en la población y en los grupos de toma de decisión” [24]. Continúan “que los pronósticos del tiempo y clima son

realizados a fin de proteger la vida y la propiedad de las personas, salvar las cosechas de posibles eventos extremos, como así también la planificación de las actividades de los grupos de decisión”[24].

1.3.6. Inteligencia artificial

Durrani et al., “los recursos de energía renovable se están volviendo más populares como un reemplazo viable de los combustibles fósiles. Si bien los combustibles fósiles todavía se utilizan para generar electricidad, conducen a la contaminación y el desperdicio de gases de efecto invernadero”[25]. Prosigue *Durrani et al.* “que los recursos renovables, sin embargo, son duraderos y libres de residuos y contaminación”[25]. *Kanchev et al.*, “estas razones han llevado a que más unidades comerciales, industriales y residenciales instalen paneles solares, turbinas eólicas y otras fuentes de energía renovable, transformándolos de consumidores habituales en **microrredes inteligentes**”[26]. *Sabzehgar et al.*, “las **microrredes inteligentes** permiten a los consumidores controlar su uso de energía de una manera más flexible, confiable y económicamente eficiente”[27]. “Sin embargo, dado que las **microrredes inteligentes** dependen de recursos renovables, a veces pueden no ser confiables debido a su dependencia de parámetros estocásticos como el clima, la energía del viento y la irradiancia solar”[27].

En otras palabras, “los recursos de generación de energía renovable como la eólica y la solar se comportan de manera estocástica debido a los cambios en los parámetros climáticos. Además, las pérdidas de transmisión y distribución influyen en el rendimiento de estos recursos”[27].

1.3.7. Aprendizaje Automático

Riabani et al., “es una rama de la Inteligencia Artificial que busca que las computadoras sean capaces de aprender e inducir conocimiento de datos. El aprendizaje automático se clasifica según el tipo de aprendizaje Supervisado y no supervisado, figura 1”[28]. Además “que el aprendizaje

supervisado está compuesto por dos fases o etapas la primera conocida como **Entrenamiento** donde se observa las características más relevantes de un evento que actúan como datos de entrada”[28] y finalmente “la segunda conocida como **Prueba** en la que compara estas características con nuevos datos de prueba y se obtiene al final de estas dos etapas un modelo entrenado y probado que intenta explicar los datos de entrada, fig.2”[28].

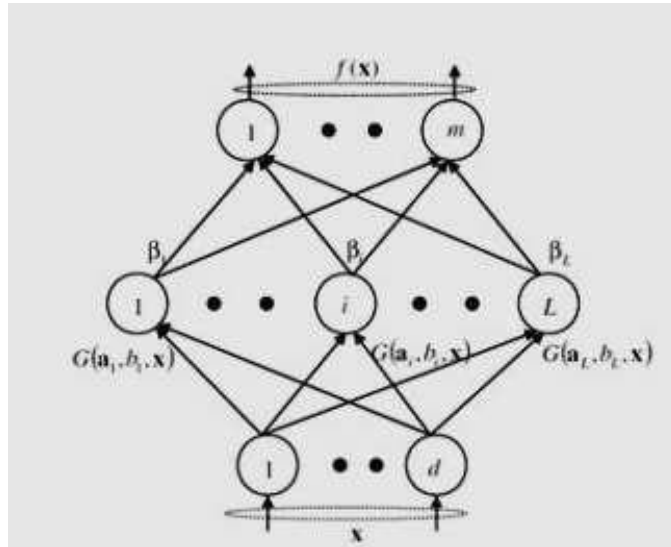


Figura 2 “Aquatinter Single Layer Feedforward neural network (SLFN)”[28]

Radiación Solar (Rs)

Mohammadi et al., “es la principal fuente de energía para todos los procesos bioquímicos y físicos en la superficie de la Tierra y la vida humana depende de ella”[29]. “Rs se considera un componente esencial de los modelos globales (circulación atmosférica general), regionales (ecológicos) y de pequeña escala (crecimiento de cultivos)”[29]. “Es uno de los parámetros críticos en generación de energía renovable refiere **Uckan, and Khudhur**”[30], “la hidrología y gestión de recursos hídricos, **Zhang et al.**[31], “los modelos de balance hídrico, **Olanike and Chandra**”[32], “los estudios climáticos, **Jahani and Mohammadi**”[33], “la urbanización, Liu et al.”[34] y “los modelos de simulación de crecimiento vegetal **Yang et al.**”[35]. Además, **Wang et al.**, “que la Rs es una parte integral de varios modelos climáticos, por

ejemplo, en el pronóstico del tiempo, el cambio climático, la circulación atmosférica general y el modelado de ecosistemas”[36]. Por lo tanto, *Liu et al.*, “muchos estudios se han centrado en la estimación precisa y el uso de patrones de radiación para diferentes propósitos[37]. Considera, *Mohammadi and Mehdizadeh* “que la intensidad y las variaciones de R_s se han utilizado en muchas áreas de estudio”[38], *Yaniktepe et al.*, “como el diseño arquitectónico, la teledetección, el modelado del crecimiento de las plantas, los sistemas de energía solar, la meteorología agrícola, la estimación de la evapotranspiración, la evaluación de las necesidades de agua de los cultivos y la gestión de los recursos hídricos”[39].

1.3.8. Clima

Según, *Laqui* “es un conjunto de elementos y factores atmosféricos, es de suma importancia en la vida cotidiana y en las actividades industriales, agrícolas, ganaderas, etc. Su estudio es para tener conocimiento del mismo y facilitar las actividades de la vida cotidiana”[40]. “El clima sus elementos y factores atmosféricos se relacionan entre ellos para mantener un equilibrio en la actividad climática como: temperatura, precipitaciones, estaciones del año entre otras más. El incremento demográfico del hombre ha alterado ciertos factores, perjudicando el equilibrio climático”[40].

1.4. Realidad problemática

En el siglo XXI, la Tierra sigue su curso a lo largo de la línea temporal, experimentando diversos eventos cósmicos y astronómicos. Estos fenómenos han llevado a expertos de todo el mundo a comprometerse en la investigación y exploración de eventos relacionados con el clima, el tiempo, los recursos hídricos, los océanos y el medio ambiente. Su objetivo es contribuir a que las naciones avancen hacia la implementación de la “Agenda 2030 de las Naciones Unidas y los Objetivos de Desarrollo Sostenible” (ODS) desempeñan un papel esencial en la orientación de las acciones y políticas para abordar desafíos globales y promover un desarrollo sostenible.

Los expertos de distintos fragmentos del mundo reconocen explícitamente que los 17 ODS y la hoja de ruta para el año 2030 representan un marco para la formulación de proyectos únicos e integrados que se alinean con la realidad del planeta Tierra. Los expertos subrayan que la naturaleza interconectada de los ODS es fundamental y permite que organizaciones como la “Organización Meteorológica Mundial (OMM) y el Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI)” tengan la responsabilidad de operar, supervisar, coordinar y mantener una extensa “Red Nacional de más de 900 Estaciones Meteorológicas e Hidrológicas”.

En los últimos años, el país en general ha experimentado un aumento significativo en las temperaturas y en la “emisión de gases de efecto invernadero”, alcanzando niveles preocupantes. Este crecimiento gradual está en gran parte relacionado con el cambio climático es una preocupación cada vez mayor, los patrones climáticos variables, la hidrología, los eventos meteorológicos extremos, las inundaciones y las sequías. Estos fenómenos reflejan el reconocimiento de que los desastres naturales pueden socavar la estabilidad a largo plazo de las sociedades y las economías, a menudo obstaculizando el progreso hacia el desarrollo sostenible durante muchos años.

En la situación actual de la atmósfera en una ubicación y momento específicos en la región de Ica, la utilización de instrumentos de campo meteorológico posibilita la observación y el seguimiento de diversos elementos meteorológicos, que comprenden la precipitación, la temperatura, la velocidad y dirección del viento, así como la humedad[11]. El clima tiene la capacidad de cambiar en un lapso corto de tiempo, incluso en cuestión de horas y días. En tiempos pasados, los patrones climáticos solían ser predecibles de manera bastante fiable a través del conocimiento indígena. Sin embargo, estos métodos tradicionales de predicción del clima han perdido su fiabilidad en la actualidad [41].

Con esta realidad y el crecimiento acelerado de la Inteligencia Artificial, nos hicimos la interrogante:

1.4.1. Problema general

¿"De qué manera el modelamiento de los parámetros meteorológicos mediante inteligencia artificial contribuye significativamente en el clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022"?

1.4.2. Problemas específicos

PE1 : ¿Como los elementos meteorológicos mediante la inteligencia artificial contribuye significativamente en el "clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022"?

PE2 : ¿De qué manera los parámetros meteorológicos mediante inteligencia artificial hacen posible prevenir significativamente los fenómenos atmosféricos en el "Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022"?

1.5. Objetivos de la investigación

1.5.1. Objetivo general

"Establecer el modelamiento de los parámetros meteorológicos que mediante inteligencia artificial contribuye significativamente en el clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022".

1.5.2. Objetivos Específicos

OE1 : Analizar que el elemento temperatura mediante la inteligencia artificial contribuye significativamente en el "clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022".

OE2 : Predecir que los parámetros meteorológicos mediante inteligencia artificial hacen posible prevenir significativamente los fenómenos

atmosféricos en el “Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022”.

1.6. Hipótesis de la investigación

1.6.1. Hipótesis general

“El modelamiento de los parámetros meteorológicos mediante inteligencia artificial contribuye significativamente en el clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022”.

1.6.2. Hipótesis Específicas

HE1: El elemento temperatura mediante la inteligencia artificial contribuye significativamente en el clima en el “Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022”.

HE2: Los parámetros meteorológicos mediante inteligencia artificial hacen posible prevenir significativamente los fenómenos atmosféricos en el “Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022”.

1.6.3. Variables

Variable independiente

Modelamiento en la estación meteorológica: El modelamiento en una estación meteorológica abarca el empleo de técnicas estadísticas y modelos matemáticos para analizar y predecir la variabilidad de los parámetros

Variable dependiente

Inteligencia artificial: La IA es el campo de la informática que se ocupa del desarrollo de sistemas capaces de aprender y adaptarse, y de realizar tareas complejas como el reconocimiento de patrones, la toma de decisiones y el procesamiento del lenguaje natural. La IA puede ser aplicada en diversas áreas, desde asistentes virtuales y sistemas de recomendación hasta vehículos

autónomos y diagnósticos médicos, contribuyendo así a la automatización y mejora de procesos en distintos sectores.

1.7. Justificación e importancia de la investigación

1.7.1. Justificación

La ciudad de Ica se encuentra en una ubicación estratégica a nivel regional, caracterizada por contar con condiciones climáticas ideales para llevar a cabo investigaciones relacionadas con la actividad solar y el clima espacial. Esto se ha convertido en un componente crucial del creciente cambio climático, que está siendo impulsado por un crecimiento poblacional desordenado y acelerado.

En particular, la exposición a niveles elevados de radiación ultravioleta es una preocupación constante en la región, superando los límites establecidos por la “Organización Mundial de la Salud (OMS) y la Organización Meteorológica Mundial (OMM)”. Esto ha dado lugar a problemas de salud, como un aumento en los casos de cáncer de piel y ceguera, especialmente entre los trabajadores del sector de las agroexportadoras, que suman más de aproximadamente 70,000 personas. La falta de una adecuada protección para esta población es una cuestión que merece atención y acción. El presente estudio de investigación se justifica por la necesidad de recopilar datos meteorológicos. La torre meteorológica es la herramienta más adecuada para esta tarea, ya que permite medir una amplia gama de variables meteorológicas que se encuentra instalado en el “Centro de Investigación para el Estudio de la Actividad Solar y sus Efectos sobre la Tierra (CIEASCEST)” es una institución dedicada a la investigación de la actividad solar y su influencia en la Tierra. En este centro, se lleva a cabo el registro de diversos elementos meteorológicos, que incluyen la temperatura ambiental, la humedad relativa, la velocidad del viento, la presión atmosférica y la radiación solar. Estos datos son esenciales para comprender y analizar el comportamiento de la atmósfera y su relación con la actividad solar [41].

De acuerdo con la esencia del proyecto mencionado, se resalta su gran importancia. El objetivo principal es realizar predicciones basadas en modelos generados en la estación meteorológica. Esto permitirá que la inteligencia artificial desempeñe un papel significativo en la gestión del clima. El propósito final es contribuir a la reducción de desastres naturales y concienciar a la población de la región de Ica sobre la necesidad de utilizar de manera responsable las variables meteorológicas.

Este proyecto tiene el potencial de posicionar a nuestra institución universitaria como un líder a nivel regional y nacional en la gestión de datos y variables climáticas.

1.7.2. Importancia

La importancia de la investigación desarrollada en el "Modelo de Predicción Climatológica con Inteligencia Artificial (AI) en la Región Ica, 2022" es de gran relevancia tanto a nivel regional como nacional. En primer lugar, esta investigación en el uso de la torre meteorológica representa un avance significativo en el campo de la predicción climatológica, al emplear tecnologías de Inteligencia Artificial para mejorar la precisión y la anticipación de los eventos meteorológicos en una región geográfica que históricamente ha experimentado condiciones climáticas desafiantes.

La aplicación de la Inteligencia Artificial en la predicción climática tiene un impacto directo en la mitigación de desastres naturales. Al permitir una mayor capacidad de pronóstico, se pueden tomar medidas preventivas con anticipación para proteger a la población y los recursos de la región. Esto incluye la capacidad de prever inundaciones, sequías, tormentas y otros fenómenos meteorológicos extremos que pueden tener un impacto devastador en la vida de las personas y la economía.

Además, esta investigación tiene un componente educativo y de concienciación fundamental. Al contribuir a una comprensión más profunda del clima y sus variaciones en la Región Ica, se brinda la oportunidad es

fundamental instruir a la población local acerca de la relevancia de la gestión sostenible de los recursos naturales y la adaptación al cambio climático para fomentar una mayor conciencia y participación en estas prácticas. El conocimiento generado a partir de este proyecto puede ser compartido con la comunidad, fomentando la toma de decisiones informadas y la implementación de prácticas más sostenibles y respetuosas con el medio ambiente es fundamental para promover la conservación de los recursos naturales y contribuir a la mitigación de los impactos negativos del cambio climático.

Por último, el hecho de que esta investigación coloque a la institución universitaria en una posición de liderazgo en el manejo de datos climáticos no solo resalta su excelencia académica, sino que también la convierte en un referente regional y nacional en la gestión de variables climáticas. Esto puede abrir puertas a colaboraciones con otras instituciones, fomentando la investigación interdisciplinaria y el intercambio de conocimientos para abordar desafíos climáticos comunes en el país.

En resumen, el "Modelo de Predicción Climatológica con Inteligencia Artificial (AI) en la Región Ica, 2022" es una investigación de gran importancia debido a su impacto en la seguridad, la educación, la concienciación y el liderazgo regional en el ámbito de la climatología. Su contribución a la mitigación de desastres naturales y al desarrollo sostenible en la región no puede subestimarse.

II. ESTRATEGIA METODOLÓGICA

La estrategia metodológica es la guía que nos permitirá elegir las técnicas, métodos y procedimientos adecuados para responder a los criterios que nos planteamos en nuestra investigación.

2.1. Área de estudio

“Se localiza en el Provincia y Región de Ica, Santiago es el distrito con mayor territorio de toda la provincia de Ica, cuenta con una población de 27,645 habitantes (según Censo INEI 2017), tiene una altitud 378 m.s.n.m.” (según Google Earth Pro).



Figura 3 Ciudad de Ica. Departamento de Ica

“Ica es una de las veinticuatro divisiones políticas del Perú, situada en la región centro-oeste del país. Limita al norte con Lima, al este con Huancavelica y Ayacucho, al sur con Arequipa y al oeste con el Océano Pacífico”. Es conocido por su diversidad geográfica, que incluye tanto áreas costeras como sierras y valles fértiles. Ica es reconocida por su producción de vinos y piscos, así como por sus hermosos paisajes

desérticos y playas. El departamento es la ciudad de Ica, que es un importante centro económico y cultural de la región [42].



Figura 4 Área donde se encuentra la estación meteorológica, ubicada en el campus de la Universidad Nacional “San Luis Gonzaga” (Av. los maestros S/N)

2.2. Tipo, nivel y diseño de investigación

2.2.1. Tipo de investigación

Investigación Aplicada con un enfoque cuantitativo. se buscó aplicar el conocimiento existente para resolver los problemas específicos. Además, el enfoque cuantitativo implicó el uso de datos numéricos y métodos cuantitativos para recopilar y analizar información en el proceso de estudio[43].

2.2.2. Nivel de investigación

Implica el análisis de series temporales para datos meteorológicos y la predicción de valores futuros mediante la aplicación de inteligencia artificial, ubicándose en el nivel de investigación "Predictivo" [43].

2.2.3. Diseño de la investigación

Es un enfoque observacional-prospectivo, lo que implica la observación de fenómenos pasados y futuros, la recopilación de datos a lo largo del tiempo y la manipulación de variables para estudiar sus efectos, lo que proporciona un enfoque integral para abordar las preguntas de investigación en el estudio [43].

2.3. Población y muestra

2.3.1. Población

La población de estudio en este caso estuvo compuesta por todos los registros de datos de las variables meteorológicas recopilados en el Centro de Investigación Solar desde junio de 2019 hasta junio de 2022. La unidad de análisis del universo corresponde a cada uno de estos registros de datos, que fueron objeto de estudio en la investigación.

2.3.2. Muestra

La muestra de investigación se determinó a través del análisis de los componentes de la serie de tiempo, con el objetivo de identificar la tendencia, ciclos, estacionalidad y elementos aleatorios presentes en los datos. Este proceso ayudó a seleccionar los datos de prueba que se utilizaron en la

investigación, lo que permitió un enfoque más específico y significativo en el análisis de las variables meteorológicas.

2.4. Técnicas de recolección de datos

Observación de campo: Los datos fueron recopilados por los equipos de la estación meteorológica, que incluyen observaciones de campo y equipos de medición de variables meteorológicas. Estos equipos capturan una serie de variables climáticas esenciales para el estudio:

1. Presión atmosférica
2. Temperatura máxima y mínima del ambiente
3. Temperatura promedio del ambiente
4. Humedad relativa
5. Radiación
6. Velocidad del viento (a los 2 m. y a los 10 m.)
7. Dirección del viento (a los 2 m. y a los 10 m.)

Los datos se extrajeron y se guardaron en un archivo de **tipo csv** llamado "**clima.csv**" y se procesarán utilizando técnicas y herramientas de Inteligencia Artificial para su análisis.

2.5. Instrumentos de recolección de datos

Mediante los instrumentos y sensores instalados en la torre meteorológica del Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, se recopilan datos meteorológicos. Estos instrumentos son fundamentales para la obtención de información precisa sobre las condiciones meteorológicas en la región:

1. Termohigrómetro (mide la temperatura ambiental-°C y la humedad relativa-%)
2. Piranómetro (mide la radiación solar directa y difusa-watts/m²)
3. Anemómetro (mide la velocidad del viento-m/s)
4. Barómetro (mide la presión atmosférica-mbar).

2.6. Técnicas de procesamiento de datos

Según *Weng et al.*, se utilizarán técnicas de procesamiento de datos meteorológicos basadas en Inteligencia Artificial. Para alcanzar los objetivos de la investigación, se empleó el software Orange Data Mining, que ofrece una amplia gama de herramientas y algoritmos de inteligencia artificial para el análisis de datos. El desarrollo de la investigación se llevó a cabo siguiendo la metodología CRISP-DM (Cross-Industry Standard Process for Data Mining), que proporciona un marco estructurado para la minería de datos. Estas técnicas y herramientas son esenciales para el procesamiento y análisis de los datos meteorológicos recopilados en el estudio [44]

2.7. Análisis e interpretación de datos

Según *Weng et al.*, para el análisis e interpretación de la tendencia, ciclos, estacionalidad y aleatoriedad en los datos meteorológicos, se utiliza el software Anaconda y la herramienta Jupyter Notebook, que se basan en el lenguaje de programación científico Python y hacen uso de librerías especializadas en el tratamiento de series de tiempo. El modelo de red neuronal se utiliza para analizar la predicción de las variables meteorológicas, previa validación del modelo. Los resultados se evalúan mediante la precisión del modelo, que se acerca a la unidad, o a través de la métrica del error cuadrático medio (MSE/Mean Square Error), que se mantiene cercano a cero. Estas técnicas permiten un análisis y una interpretación precisos de los datos meteorológicos [44].

III. RESULTADOS

3.1. Establecer el “modelamiento de los parámetros meteorológicos y la inteligencia artificial en el clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022”.

1. Comprensión del problema, “mediante la metodología CRISP-DM (Cross-Industry Standard Process for Data Mining)”.

Las variables meteorológicas en nuestra universidad se están registrando desde el año 2019 en el CIEASEST a través de la torre meteorológica, información que se registra año a año, sin tener ningún objetivo específico por lo que no es aprovechado para alguna actividad, a pesar de que en nuestra región es de suma importancia debido a la posibilidad de aprovecharse para beneficio de la enorme cantidad de agroexportadoras (más de 300) y que sería un buen aporte para conocer cómo se mueven estas variables que inciden directamente en los cultivos como la temperatura, la radiación, la velocidad y dirección del viento y otras. Igualmente, con relación al medio ambiente es muy importante para muchas actividades de la vida humana. El cambio climático está alterando la forma en que vivimos como también en el mundo. En ese escenario es importante aprovechar la información recogida por los sensores del CIEASEST. La investigación busca aprovechar los datos para realizar predicciones sobre el comportamiento de estas variables de interés: la Temperatura, la Radiación, la velocidad y dirección del viento. Estas variables meteorológicas son de mucha importancia por el enorme potencial que tienen y que se está empleando para la generación de energías no renovables de forma muy eficiente y con muy buenos resultados en nuestra región implementadas en las zonas de Ocucaje y Marcona. como lo demuestra el paso de enormes camiones con algunos equipos (aspas de molinos) para esta actividad. Es importante resaltar que en el país se genera solo 5% de energía no renovables, siendo necesario que el país por lo menos genere 20% de esta energía en palabras del ex ministro de energía y minas Pedro Gamio.

Esta problemática de necesidad de contar con energía limpias en beneficio de la estabilidad del planeta no obliga a aprovechar la gran cantidad de datos existentes para la generación de energía solar (de la cual nuestra región posee una gran llegada de la radiación y temperatura), como así también para el aprovechamiento de la velocidad y dirección del viento para aprovechar la energía eólica.

2. Comprensión de los datos

Para comprender los datos, se utilizó el software Orange dataming, este software posee una serie de widgets que se utilizar para la manipulación y procesamiento de los datos. Para comprender los datos, el archivo extraído de la fuente principal con el nombre de clima.csv se carga al software por medio del widget destinado para ello.

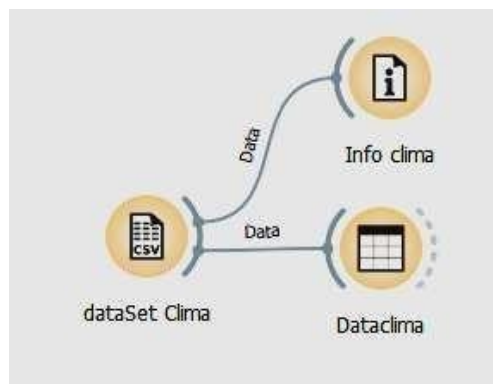


Figura 5 Modelo inicial de comprensión de los datos

En la Fig. 5, se presentan tres widgets. El primero, "**CSV import file**," se utiliza para cargar el archivo CSV llamado "**clima.csv**" El segundo widget es "**Data info**" ubicado en la parte superior, que proporciona información sobre los datos, como estadísticas descriptivas y características de las variables. Por último, el tercer widget es "**Data table**", situado en la parte inferior, que muestra los datos en formato tabular para su visualización y exploración detallada. Estos widgets son herramientas importantes para el análisis de los datos meteorológicos.

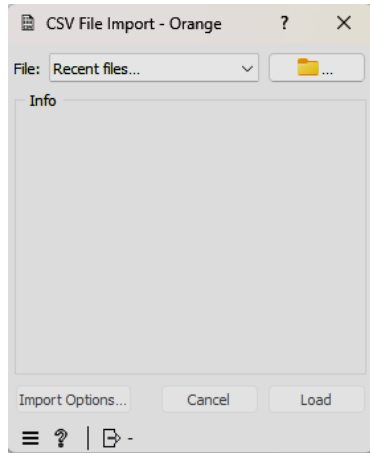


Figura 6 Carga del archivo clima.csv

En la Figura 6, se puede observar un widget que facilita la carga del archivo "**clima.csv.**" Este widget incluye un icono de carpeta que permite buscar y seleccionar el archivo específico que se desea utilizar en el análisis de datos. El icono de carpeta facilita la ubicación y selección del archivo en el proceso de carga de datos. En este contexto, el archivo que se utiliza se denomina "clima.csv" y está almacenado en la computadora, como se muestra en la Fig. 7. Este archivo es esencial para llevar a cabo el análisis de datos meteorológicos en el estudio.

Nombre	Fecha de modificación	Tipo	Tamaño
00.0_Proyecto Tesis Doctorado_Quispe G...	16/06/2023 20:09	Documento de Mi...	672 KB
Analisis variables meteorologica	09/10/2023 7:35	Orange Workflow	23 KB
Clima	03/03/2022 20:07	Archivo de valores...	751.376 KB
Data_H	05/08/2022 9:54	Archivo de valores...	2.788 KB
Informe Final - Tesis	09/10/2023 8:43	Documento de Mi...	1.979 KB

Figura 7 Selección del archivo clima.csv

Después de cargar el archivo, es necesario asignar un tipo de datos a cada una de las columnas. Para lograr esto, se utiliza la opción "**Column type**" y se elige el tipo de dato correspondiente, como se muestra en la Fig. 8. Por ejemplo, la columna de "Fecha" se asigna como tipo "**Datetime.**" Este tratamiento se reitera para cada una de las columnas en el archivo. Además, se tiene la flexibilidad de

omitir o ignorar columnas que no son relevantes o necesarias para el análisis. Este paso es esencial para garantizar que los datos se interpreten correctamente y se ajusten a los requisitos de la investigación.

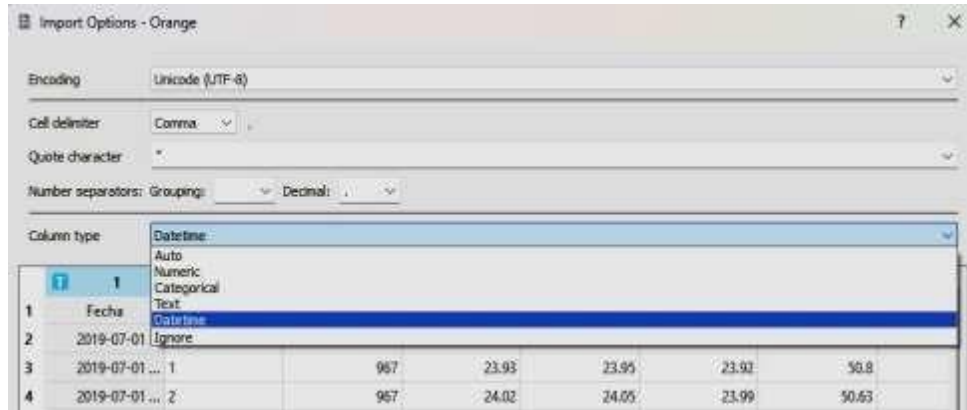


Figura 8 Asignación del tipo de dato a las columnas

Column	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Fecha	registros	Frecuen	Presig	Tras	Tras	Hvnt	Radiacion	Volumen2	Disentor2	Volumen18	RadiacionT	Disent
2	2019-07-01 ... 0		944	23.78	23.82	23.65	50.77	418.2	0.09	81.27	0.334	0.04382163	
3	2019-07-01 ... 1		967	23.93	23.95	23.82	50.8	407	0	334.3	0.075	0.004369996	
4	2019-07-01 ... 2		967	24.02	24.05	23.99	50.63	418.8	1.1	228.1	1.23	0.004367238	
5	2019-07-01 ... 3		967	23.99	24.02	23.95	50.59	415.8	0.95	306.1	1.35	0.00437795	
6	2019-07-01 ... 4		967	24	24.02	23.99	50.74	415.3	1.175	301.4	1.325	0.004353023	
7	2019-07-01 ... 5		967	24	24.02	23.99	50.67	415	1.175	312	0.95	0.004349617	
8	2019-07-01 ... 6		967	23.99	24.02	23.95	50.77	414.3	0.8	306.5	0.975	0.004342819	
9	2019-07-01 ... 7		967	23.99	23.99	23.99	50.84	411.7	0.925	318.8	0.315	0.004317177	
10	2019-07-01 ... 8		967	24	24.02	23.99	50.84	412.9	0	318.5	0.175	0.004326254	
11	2019-07-01 ... 9		967	23.99	23.99	23.99	50.84	412.1	0	305.7	0.65	0.004321311	
12	2019-07-01 ... 10		967	23.83	23.95	23.82	50.9	411.8	0.25	258.9	1.025	0.004315672	
13	2019-07-01 ... 11		967	23.85	23.99	23.82	50.87	411.8	0.95	306.1	0.25	0.004315672	
14	2019-07-01 ... 12		967	23.95	23.99	23.82	50.9	410.5	0.25	292	0.25	0.004305454	
15	2019-07-01 ... 13		967	23.97	23.99	23.95	50.74	410.1	0	251.5	0	0.004300067	
16	2019-07-01 ... 14		967	23.87	23.99	23.85	50.7	418.3	0	208.8	0.725	0.004292984	

Figura 9 Tabla con los datos y sus tipos de datos asignados

En la Fig. 9, una vez que se ha cargado el archivo y se ha completado la asignación de los tipos de datos correspondientes a cada columna, se puede observar que en la parte superior de cada columna se muestra el tipo de dato asignado. Finalmente, se confirma la operación seleccionando la opción "OK," lo que significa que los datos han sido cargados y configurados correctamente en

el software. Este paso finaliza el proceso de preparación de los datos para su análisis en la herramienta.

A continuación, se procedió a la visualización de los datos. Antes de esto, se utilizó el widget "Data info" para obtener una vista general de los datos, como se muestra en la Fig. 10. Esta herramienta proporciona información general sobre los datos antes de proceder a su visualización y análisis en detalle.

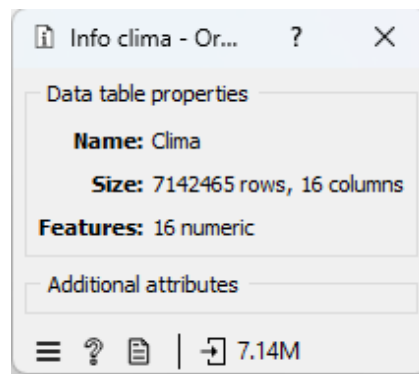


Figura 10 Information del dataSet Clima

De acuerdo a la información proporcionada por el widget, el archivo de datos contiene 16 columnas de tipo numérico y un total de 7,142,465 registros de datos. Esto brinda una visión general de la estructura y la magnitud del conjunto de datos que será objeto de análisis en la investigación.

	Fecha	registro	Presion	Ftemp	Tmae	Tmae	Tmae	Hrelat	Radiacion	Velocidad
1	2019-07-01 15:37:00	0	944	23.78	23.92	23.65	50.77	438.2	1	
2	2019-07-01 15:37:10	1	967	23.89	23.95	23.92	50.8	437	1	
3	2019-07-01 15:37:20	2	967	24.02	24.05	23.99	50.63	436.8	1	
4	2019-07-01 15:37:30	3	967	23.98	24.02	23.95	50.59	435.8	1	
5	2019-07-01 15:37:40	4	967	24	24.02	23.99	50.74	435.3	1	
6	2019-07-01 15:37:50	5	967	24	24.02	23.99	50.67	435	1	
7	2019-07-01 15:38:00	6	967	23.99	24.02	23.95	50.77	434.3	1	
8	2019-07-01 15:38:10	7	967	23.99	23.99	23.99	50.84	433.7	1	
9	2019-07-01 15:38:20	8	967	24	24.02	23.99	50.94	432.9	1	
10	2019-07-01 15:38:30	9	967	23.99	23.99	23.99	50.84	432.1	1	
11	2019-07-01 15:38:40	10	967	23.93	23.95	23.92	50.9	431.6	1	
12	2019-07-01 15:38:50	11	967	23.95	23.99	23.92	50.87	431.6	1	
13	2019-07-01 15:39:00	12	967	23.95	23.99	23.92	50.9	430.5	1	
14	2019-07-01 15:39:10	13	967	23.97	23.99	23.95	50.74	430.1	1	
15	2019-07-01 15:39:20	14	967	23.97	23.99	23.95	50.7	429.3	1	
16	2019-07-01 15:39:30	15	967	23.97	23.99	23.95	50.67	428.7	1	
17	2019-07-01 15:39:40	16	967	24	24.02	23.99	50.6	427.6	1	
18	2019-07-01 15:39:50	17	967	24.02	24.02	24.02	51.11	427.3	0	
19	2019-07-01 15:40:00	18	967	24.04	24.09	23.99	50.6	427	1	
20	2019-07-01 15:40:10	19	967	24.02	24.05	23.99	50.7	426.7	1	

Figura 11 Datos del clima

Como se puede observar en la Fig. 11, mediante el uso del widget "**Data table**," se presenta la información detallada de los datos almacenados. Es importante destacar que no se han identificado datos vacíos, extraviados o perdidos (no missing data) en el conjunto de datos. Este widget permite explorar y navegar a través de todos los registros de datos de manera similar a una "hoja de cálculo en Excel", lo que facilita la inspección y comprensión de los datos en su totalidad.

Es relevante destacar que el registro de los datos se efectúa cada 10 segundos, como se puede observar en la columna de fecha. Este intervalo de tiempo entre registros es un detalle importante, ya que influye en el tratamiento y análisis apropiado de los datos. La frecuencia de muestreo, en este caso, de 10 segundos, tiene implicaciones en la interpretación y la precisión de los rendimientos en el entorno del estudio.

Comprensión gráfica de los datos

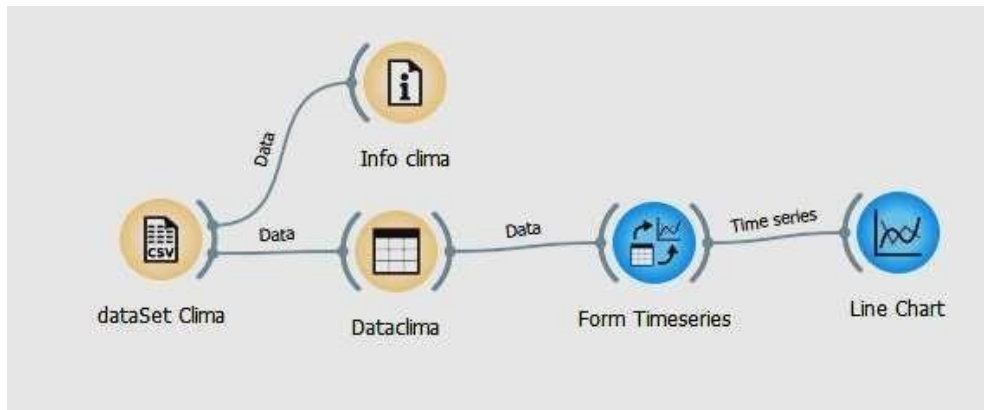


Figura 12 Modelo para comprensión gráfica

En la Fig. 12, se presenta el modelo inicial que incorpora dos widgets destinados a brindar una comprensión visual de los datos. El primer widget es "**From Timeseries**," que se utiliza para representar los datos como una serie de tiempo. El segundo widget es "**Line Chart**," que permite una visualización gráfica de los datos en forma de gráfico de líneas. Estas herramientas son esenciales para el análisis y la representación visual de los datos meteorológicos en el estudio.

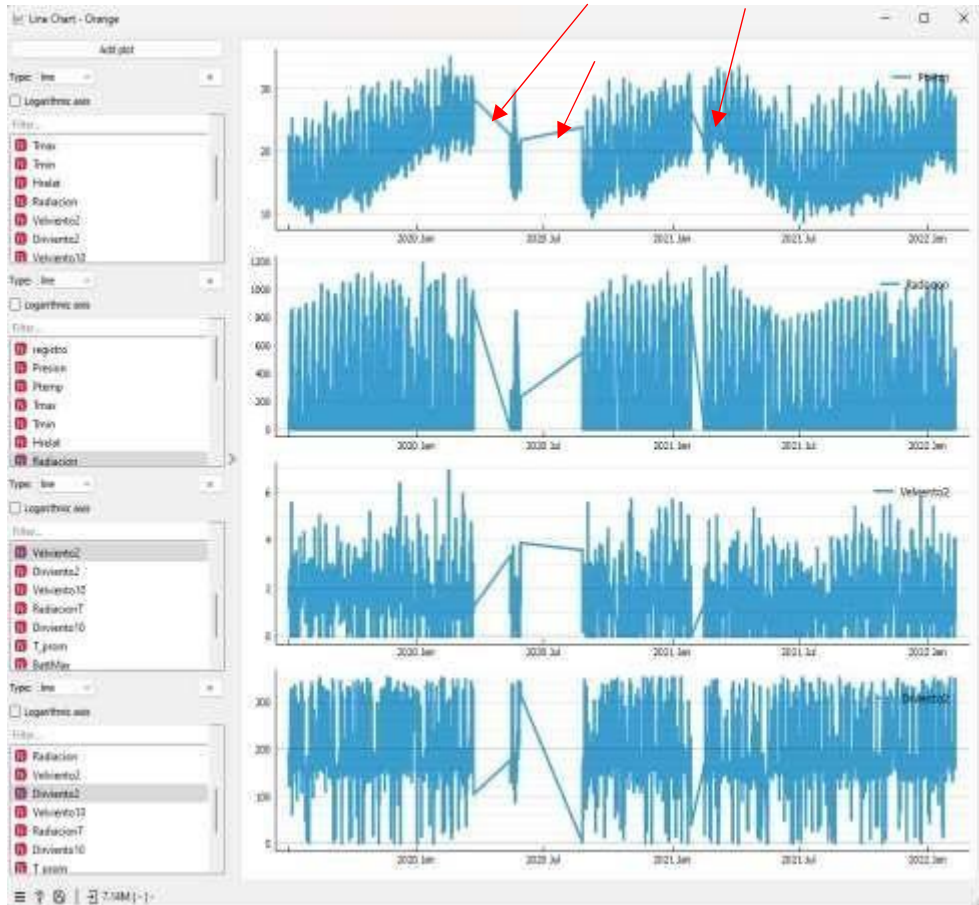


Figura 13 Comprensión gráfica de los datos

En la Fig. 13, se puede observar en las gráficas generadas con los widgets "**From Timeseries**" y "**Line Chart**" que existen espacios vacíos o huecos en los datos, los cuales se indican con flechas de color rojo. Estos huecos en los datos deben ser corregidos antes de aplicar los modelos de análisis y predicción, ya que pueden afectar la calidad y la precisión de los resultados. Es importante abordar estos datos faltantes para asegurar la integridad de los datos y la confiabilidad de los modelos.

Por lo tanto: Los resultados revelaron, a través de las gráficas generadas con los widgets "**From Timeseries**" y "**Line Chart**", la presencia de espacios vacíos o huecos en los datos, destacados con flechas de color rojo en la Fig. 13.

Estos vacíos en los datos fueron identificados como elementos críticos que requerían corrección antes de la aplicación de modelos de análisis y predicción.

La presencia de datos faltantes podría comprometer la calidad y precisión de los resultados obtenidos a través de los modelos. En consecuencia, se hizo hincapié en la importancia de abordar y corregir estos vacíos para garantizar la integridad de la información y mejorar la confiabilidad de los modelos utilizados en el estudio.

Este hallazgo subraya la necesidad de realizar un tratamiento riguroso de los datos antes de aplicar técnicas de inteligencia artificial para garantizar que los resultados sean robustos y representativos del comportamiento real de los parámetros meteorológicos en el Centro de Investigación. La implementación efectiva de estos modelos puede contribuir de manera significativa a una mejor comprensión y predicción del clima en la región de estudio, respaldando así el objetivo principal del proyecto.

3.2. Analizar el “elemento temperatura y la inteligencia artificial en el clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022”.

3. Preparación de los datos

Para poder realizar la preparación de los datos o preprocesamiento de estos, primero se va a realizar la disminución de la complejidad de los datos para tener un manejo más eficiente ya que los datos se encuentran en segundos (un *dataSet* muy grande) por conveniencia se van a minimizar la complejidad de los datos a la unidad de tiempo por horas.

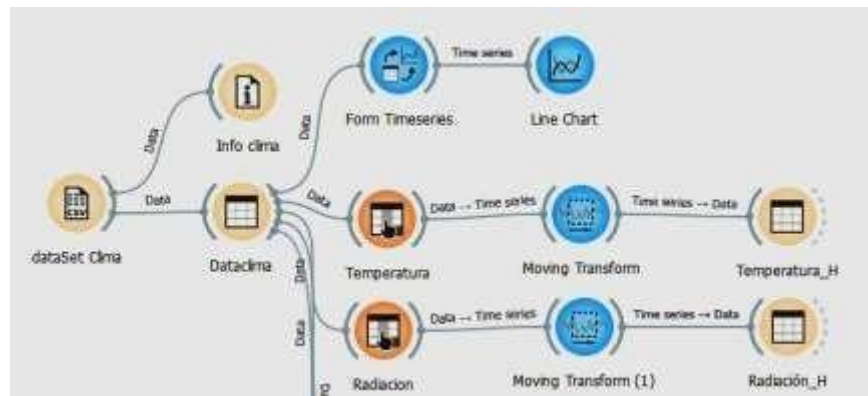


Figura 14 Transformación de las variables Temperatura y Radiación

En el caso de las variables de temperatura y radiación, se lleva a cabo una transformación utilizando los widgets "Select Columns," "Moving Transform," y "Data Table." Esta transformación tiene como objetivo convertir las unidades de dimensionalidad a unidades en horas, como se muestra en la Fig. 16. Esta operación es esencial para asegurar que las variables estén en la misma escala y unidad de tiempo, lo que facilita su análisis y comparación en el estudio.

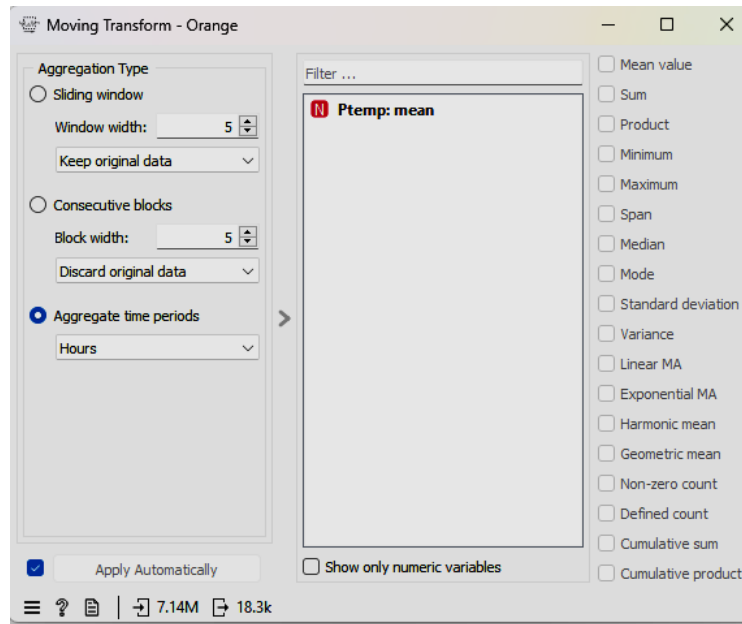


Figura 15 Transformación de la variable Temperatura

El widget "Moving Transform" se utiliza para llevar a cabo la transformación de las variables de temperatura y radiación. Para ello, se debe seleccionar el periodo de tiempo en horas (Hours), luego se elige la variable "Ptemp" y se selecciona el tipo de transformación, que en este caso es el promedio (Mean value). De manera similar, se realiza el mismo proceso para la variable de radiación.

Es importante tener en cuenta que esta transformación puede llevar tiempo, por lo que es necesario esperar a que se complete el proceso. Interrumpir la conversión antes de que termine puede afectar la calidad de los resultados y la consistencia de los datos transformados. Los resultados de la transformación se muestran en la Fig.17.

Temperatura_H - Orange			
Info			
18304 instances (no missing data)			
3 features			
No target variable.			
No meta attributes.			
Variables			
<input checked="" type="checkbox"/> Show variable labels (if present)			
<input type="checkbox"/> Visualize numeric values			
<input checked="" type="checkbox"/> Color by instance classes			
Selection			
<input checked="" type="checkbox"/> Select full rows			
	Hours	Instance count	Ptemp (mean)
1	2019-07-01 15:00:00	138	23.864
2	2019-07-01 16:00:00	464	23.3262
3	2019-07-01 17:00:00	720	19.5332
4	2019-07-01 18:00:00	720	16.471
5	2019-07-01 19:00:00	720	14.6669
6	2019-07-01 20:00:00	720	13.4795
7	2019-07-01 21:00:00	720	13.224
8	2019-07-01 22:00:00	720	13.105
9	2019-07-01 23:00:00	720	13.0419
10	2019-07-02 00:00:00	720	12.9156

Figura 16 Datos transformados de la Temperatura

En la Fig. 17, se puede verificar que se ha realizado la transformación de la variable de temperatura, convirtiéndola a una unidad de tiempo por horas, como se muestra en la columna "**Hours.**"

En el caso de las “variables de velocidad y dirección del viento”, se requiere un tratamiento adicional debido a que se cuentan con datos recopilados a dos alturas diferentes (a 2 metros y a 10 metros). Por lo tanto, el siguiente paso es combinar estos datos en una sola variable que contenga el promedio de las variaciones a estas alturas. Esto permitirá tener una representación unificada de estas variables para su posterior análisis.

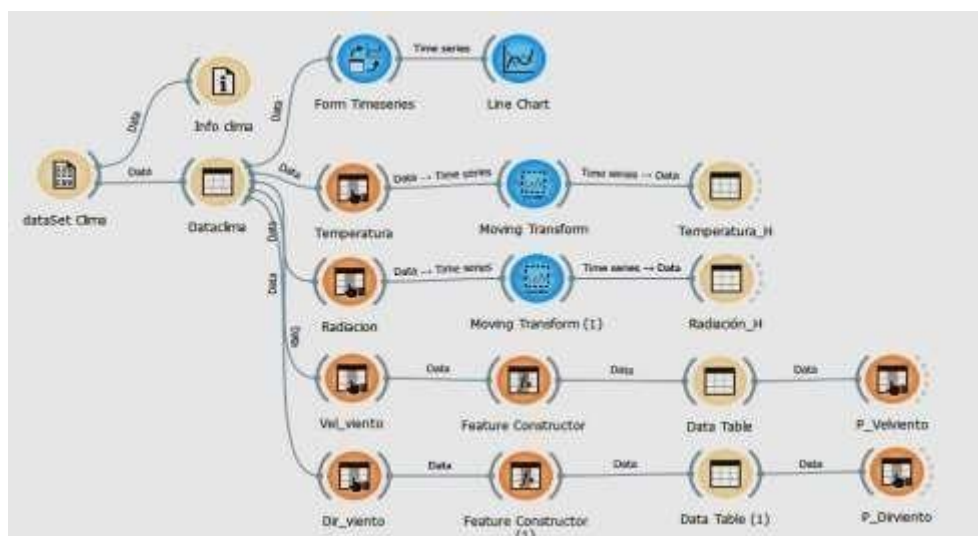


Figura 17 Ampliación del modelo con las variables velocidad y dirección del viento

En la Fig. 18, se puede observar que, para el tratamiento de las “variables de velocidad y dirección del viento”, se emplearon los widgets "**Feature Constructor**," "**Data Table**," y "**Select Column**." Los resultados de este proceso se presentan en las Figs. del 18 al 20. Estas herramientas se utilizaron para realizar la combinación de datos de las dos alturas y generar una representación unificada de estas variables, lo que simplificará su análisis en la investigación.

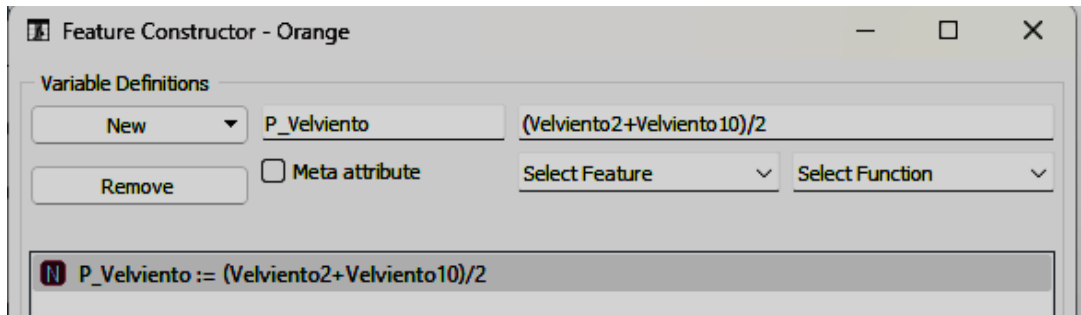


Figura 18 . Conversión de la velocidad del viento

En la Fig. 18, se lleva a cabo la creación de una fórmula para calcular el promedio de calcular la velocidad del viento utilizando los datos recopilados a 2 metros y 10 metros de altura. Para realizar esta operación, se crea un nuevo constructor de fórmulas haciendo clic en "**New**." Luego, se asigna un nombre a la nueva variable y se construye la fórmula que permite calcular el promedio de la velocidad del viento.

Esta operación es esencial para combinar los datos de las dos alturas y obtener una representación unificada de la velocidad del viento:

$$\frac{(Velviento2 + Velviento 10)}{2}$$

En el proceso de construcción de la fórmula, es necesario seleccionar las variables relevantes que se utilizarán en la fórmula. Esto se realiza a través del widget "**Select Feature**" (Selección de características). Al elegir las variables adecuadas en este paso, se asegura que la fórmula se construya de manera

correcta y que tenga en cuenta las variables necesarias para calcular el promedio de calcular la velocidad del viento utilizando los datos de las dos alturas.



	Fecha	Velviento2	Velviento10	P_Velviento
1	2019-07-01 15:3...	0.09	0.324	0.207
2	2019-07-01 15:3...	0	0.875	0.4375
3	2019-07-01 15:3...	1.1	1.25	1.175
4	2019-07-01 15:3...	0.95	1.55	1.25
5	2019-07-01 15:3...	1.175	1.325	1.25
6	2019-07-01 15:3...	1.175	0.95	1.0625
7	2019-07-01 15:3...	0.8	0.875	0.8375
8	2019-07-01 15:3...	0.325	0.325	0.325
9	2019-07-01 15:3...	0	0.175	0.0875
10	2019-07-01 15:3...	0	0.65	0.325

Figura 19 Resultados del cálculo

En el widget "**Data Table**," se puede observar que se ha agregado una nueva columna que contiene el promedio de la velocidad del viento. Además, las columnas originales que representan la velocidad del viento a 2 metros y 10 metros también se mantienen en el conjunto de datos. Esta operación permite disponer de una representación consolidada de la velocidad del viento, lo que facilita su análisis y uso en la investigación.

Para finalizar este proceso, se utiliza el widget "**Select Columns**" para seleccionar únicamente la columna que contiene la variable que se va a convertir, como se muestra en la Fig. 20. Esto asegura que solo la columna relevante sea considerada en el análisis y en los modelos de predicción, lo que simplifica la gestión de los datos y la interpretación de los resultados en la investigación.

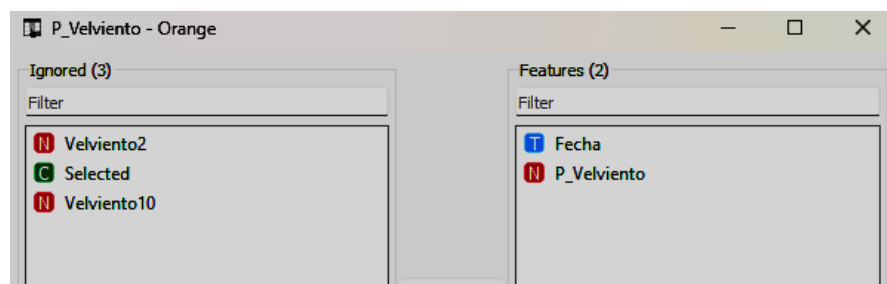


Figura 20 Selección de la variable obtenida en el cálculo

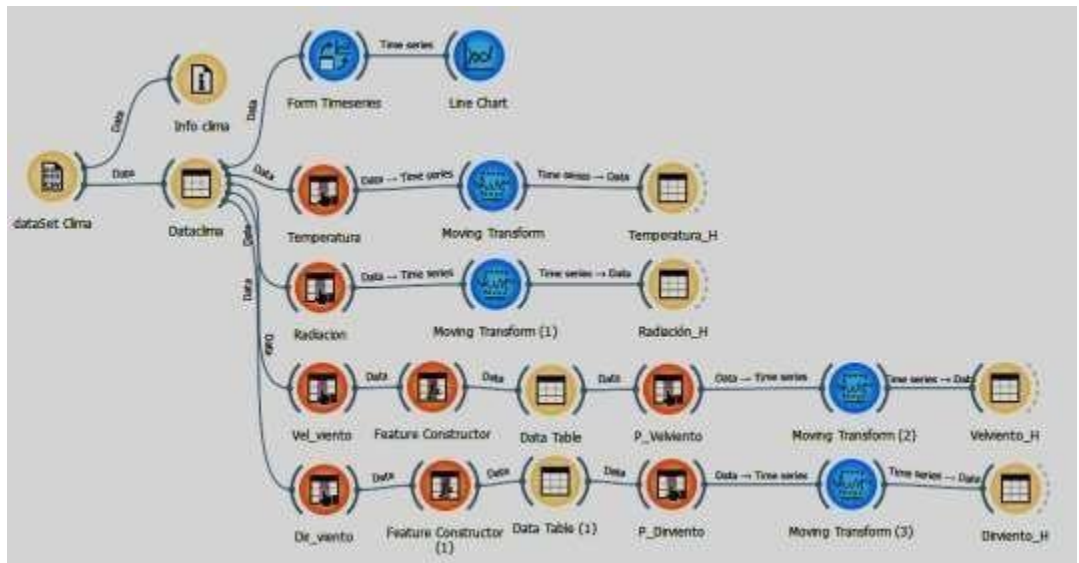


Figura 21 Modelo final de transformación de las variables meteorológicas

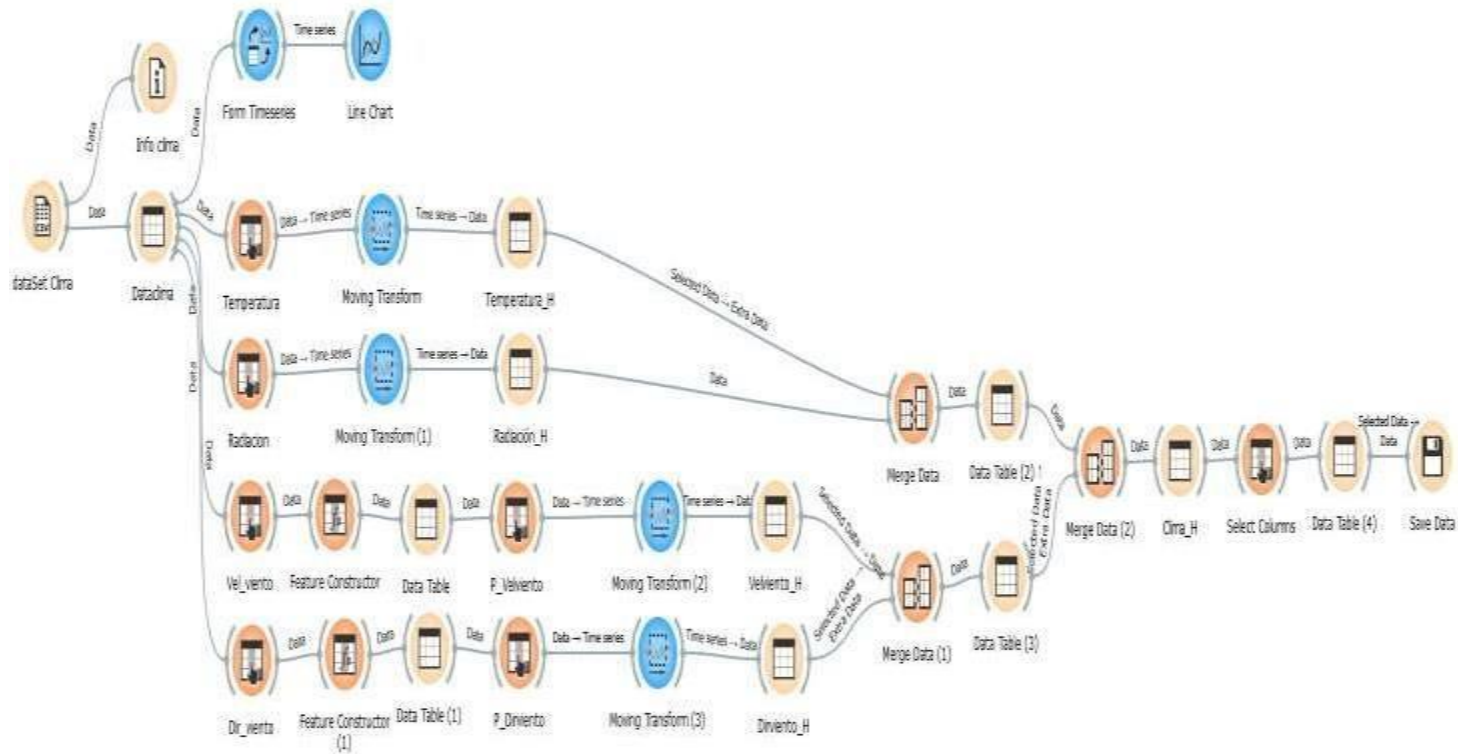


Figura 22 Obtención de la dataSet Clima_H.csv

Después de haber transformado las variables a la unidad de tiempo en horas, se procede a combinar las tablas resultantes en un solo conjunto de datos. Para lograr esto, se utilizan los widgets **"Merge Data," "Data Table," "Select Columns" y "Save Data."** Estos widgets permiten la unión de las tablas y la generación de un conjunto de datos consolidado, que en este caso se ha nombrado como "Clima_H.csv." Este proceso facilita el análisis y la gestión de los datos en la investigación.

Data Table (4) - Orange					
Info					
18304 instances (no missing data)					
5 features					
No target variable.					
No meta attributes.					
Variables					
<input checked="" type="checkbox"/> Show variable labels (if present)					
<input type="checkbox"/> Visualize numeric values					
<input checked="" type="checkbox"/> Color by instance classes					
Selection					
<input checked="" type="checkbox"/> Select full rows					
	Hours	RadiacionT (mean)	Ptemp (mean)	P_Velviento (mean)	P_Dirviento (mean)
1	2019-07-01 15:0...	0.00422589	23.864	0.734652	190.582
2	2019-07-01 16:0...	0.00181965	23.3262	1.76554	167.227
3	2019-07-01 17:0...	0.000360345	19.5332	2.95715	166.871
4	2019-07-01 18:0...	0	16.471	2.12576	159.482
5	2019-07-01 19:0...	0	14.6669	2.04222	152.386
6	2019-07-01 20:0...	0	13.4795	1.9476	166.072
7	2019-07-01 21:0...	0	13.224	2.46924	164.905
8	2019-07-01 22:0...	0	13.105	2.00705	164.625
9	2019-07-01 23:0...	0	13.0419	1.54045	168.919
10	2019-07-02 00:0...	0	12.9156	1.2417	156.575
11	2019-07-02 01:0...	0	12.7874	1.25576	157.279
12	2019-07-02 02:0...	0	12.5865	1.2684	146.556
13	2019-07-02 03:0...	0	12.653	1.29333	159.389
14	2019-07-02 04:0...	0	12.3964	1.75708	165.131
15	2019-07-02 05:0...	0	12.1153	2.16003	167.095
16	2019-07-02 06:0...	3.57238e-05	12.0361	2.12562	154.606
17	2019-07-02 07:0...	0.000367132	12.432	2.3375	154.226
18	2019-07-02 08:0...	0.000842893	12.7842	2.21038	156.12
19	2019-07-02 09:0...	0.00132537	13.5202	2.24507	150.137
20	2019-07-02 10:0...	0.00199444	14.5995	2.28174	162.29

Figura 23 datos del dataSet Clima_H.csv

En el conjunto de datos final, que se muestra en la Fig. 23, se presentan las variables meteorológicas que se han estudiado, junto con el campo de fecha expresado en horas. Este conjunto de datos proporciona una representación completa y unificada de las variables meteorológicas, lo que es esencial para llevar a cabo el análisis y la predicción en la investigación. Las variables y la información de tiempo en horas se encuentran listas para su posterior procesamiento y análisis con técnicas de inteligencia artificial.

En comparación con la visualización gráfica de los datos originales que se mostró en la Fig. 12, en la siguiente Fig. 24, se presenta el comportamiento de los datos en el tiempo, lo cual es esencial para comprender la evolución de las variables meteorológicas. Para llevar a cabo esta visualización, se creó un nuevo archivo que utiliza la información del archivo "Clima_H.csv." Este archivo se ha construido con un modelo específico diseñado para este propósito, lo que permite analizar de manera efectiva la evolución temporal de las variables meteorológicas y realizar predicciones basadas en esta información. Este análisis gráfico proporciona una representación más clara y comprensible de los datos en el contexto temporal.

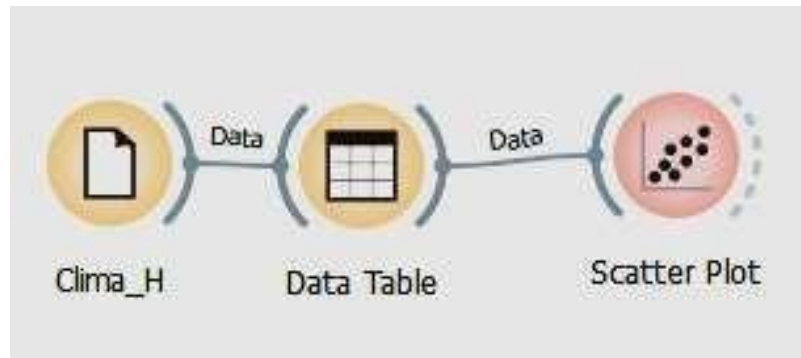


Figura 24 Modelo para análisis gráfico del archivo Clima_H.csv

Para realizar el análisis gráfico de esta reducción de la dimensionalidad temporal, se empleó un nuevo widget llamado "**Scatter Plot.**" Este widget se utiliza con el propósito de visualizar y analizar el comportamiento de las variables en relación con el tiempo. Proporciona una representación gráfica que facilita la comprensión de la relación entre las variables meteorológicas y su evolución en el tiempo.

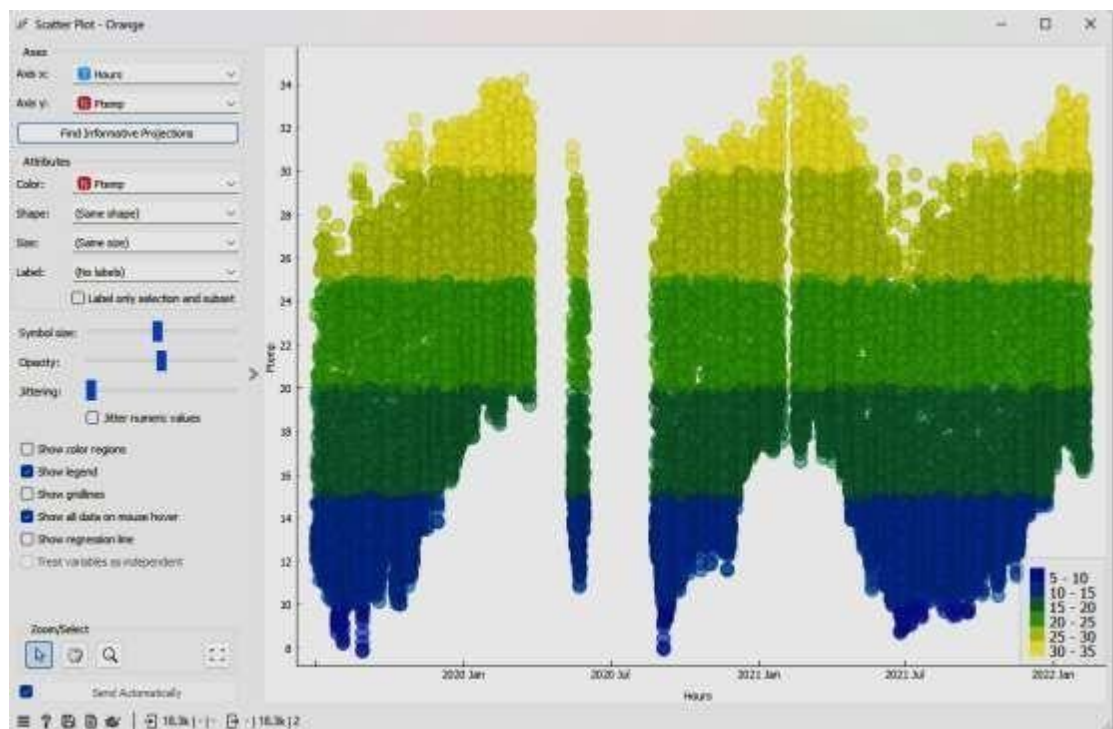


Figura 25 Gráfica del comportamiento de la variable de temperatura

Como se puede observar en el análisis gráfico de los datos expresados en unidades de tiempo en horas, la gráfica muestra la presencia de datos faltantes.

Por este motivo, es necesario completar estos datos utilizando un proceso de imputación. La identificación de los datos faltantes es fundamental para asegurar que los conjuntos de datos estén completos y permitan un análisis preciso y la efectiva aplicación de técnicas de inteligencia artificial. La imputación de datos es un proceso importante para garantizar que la información sea lo más completa y precisa posible.

La Fig. 25 muestra la gráfica de la temperatura, en la que se puede observar la periodicidad de los datos a lo largo del tiempo. Para completar los datos faltantes, se opta por utilizar datos del mismo período de tiempo para obtener un conjunto de datos final más completo. Esta estrategia de imputación de datos permite preservar la estructura temporal de los datos y garantizar que los valores faltantes se llenen de manera coherente con la tendencia observada en la serie temporal la **dataSet final**. La imputación de datos es un paso importante para asegurar que los análisis posteriores sean más precisos y significativos.

Detección de datos perdidos

Según lo representado en la Fig. 26, se ha verificado la presencia de datos faltantes, que se observan en tres periodos de tiempo específicos:

Periodo del 2020-03-23 al 2020-05-12.

Periodo del 2020-05-29 al 2020-08-23.

Periodo del 2021-01-28 al 2021-02-13.

Identificar estos intervalos con datos faltantes es esencial para llevar a cabo la imputación de datos de manera efectiva y garantizar que el conjunto de datos esté completo y listo para su análisis posterior.

Data Table (1) - Orange

Info
46 instances (no missing data)
3 features
Target with 2 values
No meta attributes

Variables
 Show variable labels (if present)
 Visualize numeric values
 Color by instance classes

Selection
 Select full rows

Selected	Hours	Radiacion	Ptemp	P_velivento	P_Diriviento	
1	No	2020-03-21 15:0...	0.00595188	31.321	4.37855	206.06
2	No	2020-03-21 16:0...	0.00358544	30.099	4.59678	305.062
3	No	2020-03-21 17:0...	0.000729324	27.9637	4.25383	309.297
4	No	2020-03-21 18:0...	2.00945e-05	26.4874	3.1241	299.863
5	No	2020-03-21 19:0...	0	25.3064	2.19047	231.091
6	No	2020-03-21 20:0...	0	23.1147	1.9767	167.111
7	No	2020-03-21 21:0...	0	22.3936	1.87056	174.716
8	No	2020-03-21 22:0...	0	22.2762	1.54087	169.022
9	No	2020-03-21 23:0...	0	22.0538	0.902153	154.165
10	No	2020-03-22 00:0...	0	21.6772	0.920486	163.153
11	No	2020-03-22 01:0...	0	21.0389	0.874479	143.156
12	No	2020-03-22 02:0...	0	20.4051	0.799479	188.054
13	No	2020-03-22 03:0...	0	19.972	0.615382	173.778
14	No	2020-03-22 04:0...	0	19.6877	0.309201	151.429
15	No	2020-03-22 05:0...	0	19.3337	0.204444	182.80
16	No	2020-03-22 06:0...	0.000254568	19.6067	0.717604	175.551
17	No	2020-03-22 07:0...	0.00234656	22.1135	1.05803	130.276
18	No	2020-03-22 08:0...	0.00479456	24.3198	1.91208	168.767
19	No	2020-03-22 09:0...	0.00593591	26.3016	1.86611	169.28
20	No	2020-03-22 10:0...	0.0087761	28.6586	1.89785	173.051
21	No	2020-03-22 11:0...	0.00951526	30.6706	1.62219	176.938
22	No	2020-03-22 12:0...	0.00965256	31.978	1.85993	150.205
23	No	2020-03-22 13:0...	0.00951985	32.9483	1.55618	142.917
24	No	2020-05-13 17:0...	8.62658e-08	22.2654	2.82857	167.276
25	No	2020-05-13 18:0...	0	21.1658	2.38529	158.536
26	No	2020-05-13 19:0...	0	19.7531	2.47628	148.846
27	No	2020-05-13 20:0...	0	18.6796	1.98174	160.626
28	No	2020-05-13 21:0...	0	17.7954	1.67562	160.678
29	No	2020-05-13 22:0...	0	17.1541	1.76764	155.093
30	No	2020-05-13 23:0...	0	16.7612	0.991424	171.196
31	No	2020-05-14 00:0...	0	16.4126	0.651111	163.403
32	No	2020-05-14 01:0...	0	16.0846	1.33469	160.823
33	No	2020-05-14 02:0...	0	15.6617	1.56007	170.017
34	No	2020-05-14 03:0...	0	15.2811	1.34444	165.591
35	No	2020-05-14 04:0...	0	14.9224	1.88479	157.284
36	No	2020-05-14 05:0...	0	15.0672	2.12873	171.584
37	No	2020-05-14 06:0...	0.000102273	15.1476	2.39146	165.737
38	No	2020-05-14 07:0...	0.00064258	15.3221	2.48690	160.357
39	No	2020-05-14 08:0...	0.00175527	16.1265	2.65184	166.54

Restore Original Order
Send Automatically

46 | 46

Figura 26 Comprobación de datos faltantes

En la Fig. 26 se verifica la presencia de datos faltantes en el primer período, que se resalta en el marco rojo.

El modelo se utilizó para realizar el proceso representado en la Fig. 27, que se ha diseñado con el propósito de identificar y manejar los datos faltantes en el conjunto de datos de manera efectiva.

La detección y manejo de estos datos faltantes son cruciales para garantizar la integridad y la calidad del conjunto de datos antes de realizar cualquier análisis o predicciones.

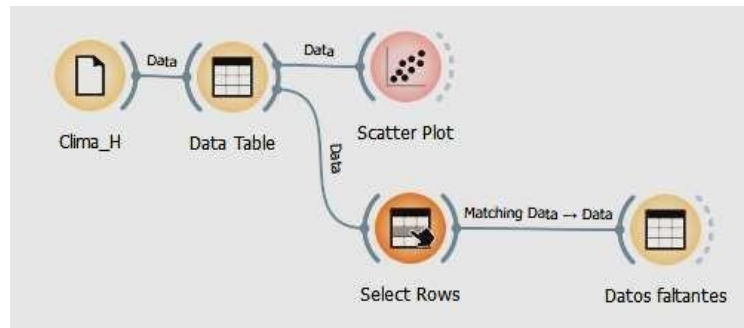


Figura 27 Verificación de los datos faltantes

Para llevar a cabo la verificación de los datos faltantes, se utilizaron los widgets *"Select Rows"* y *"Data Table"*. Se realizaron selecciones de los registros que posiblemente carecían de datos. En el widget *"Select Rows"*, se muestra el filtro respectivo, tal como se ilustra en la Figura 28. Este proceso de selección y filtrado es esencial para identificar los datos que faltan y determinar las áreas que requieren imputación o completación.

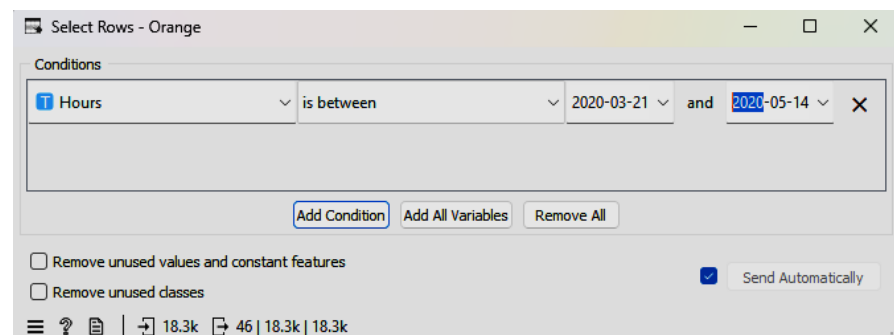


Figura 28 Comprobación de datos faltantes del primer rango de fechas

La Fig. 28 demuestra la comprobación de datos faltantes en este caso, que se ha realizado para el primer período. Para verificar los otros periodos, solo se necesita agregar una nueva condición utilizando la función "Add Condition" e incluir el rango de fechas correspondiente. Este proceso permite identificar de manera precisa y eficiente los datos faltantes en diferentes intervalos de tiempo, lo que es esencial para llevar a cabo la imputación de datos de manera adecuada.

Para completar los datos faltantes, se ha seleccionado información de los mismos periodos, pero en años anteriores o posteriores. Se ha determinado que los datos

de reemplazo para estos valores faltantes se encuentran en los siguientes periodos:

Periodo del 2021-03-23 al 2021-05-12.

Periodo del 2021-05-29 al 2021-08-23.

Periodo del 2022-01-28 al 2022-02-13.

Esta estrategia de imputación de datos ayuda a asegurar que los datos faltantes se completen de manera adecuada y consistente con la serie temporal, lo que es crucial para el análisis y las predicciones posteriores.

Para lograr la imputación de datos en base a la estrategia mencionada, se ha incorporado el modelo representado en la Fig. 29. Este modelo se ha diseñado para seleccionar y copiar los datos de los periodos adecuados en los años anteriores o posteriores y completar de manera efectiva los valores faltantes en el conjunto de datos. La imputación de datos es un paso importante en la preparación de datos antes de llevar a cabo cualquier análisis o predicciones.

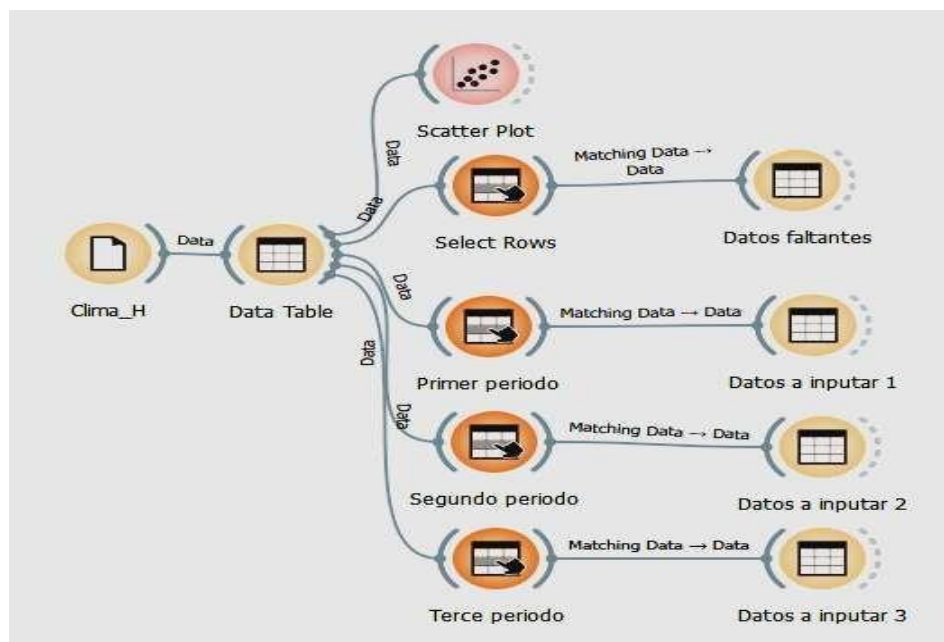


Figura 29 Selección de registro de datos a imputar

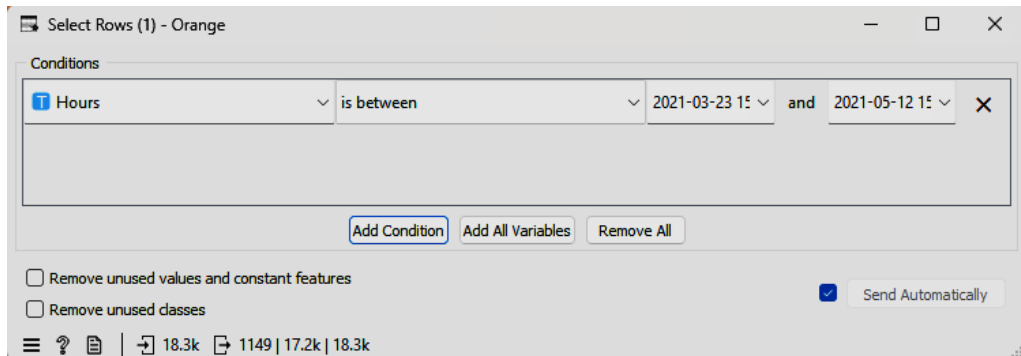


Figura 30 Datos seleccionados para el primer periodo

La Fig. 30 ilustra el proceso de filtrado de registros que se imputarán para el primer período de tiempo faltante. Este procedimiento se realizará de manera similar para los otros dos periodos de datos faltantes, asegurando así la corrección de los valores faltantes en diferentes intervalos de tiempo. La imputación de datos de esta manera garantiza la integridad y coherencia de la serie temporal, lo que es esencial para análisis y predicciones precisas.

Para consolidar todos los registros en un único archivo y crear el archivo final *Clima_H_final.csv*, se ha utilizado un nuevo widget que realiza esta tarea. Además, se ha procedido a estructurar las fechas, reduciendo el año anterior en la serie temporal. Este paso asegura que los datos estén preparados para su posterior análisis y predicciones, y que el archivo final sea coherente y consistente con la serie temporal de datos meteorológicos.

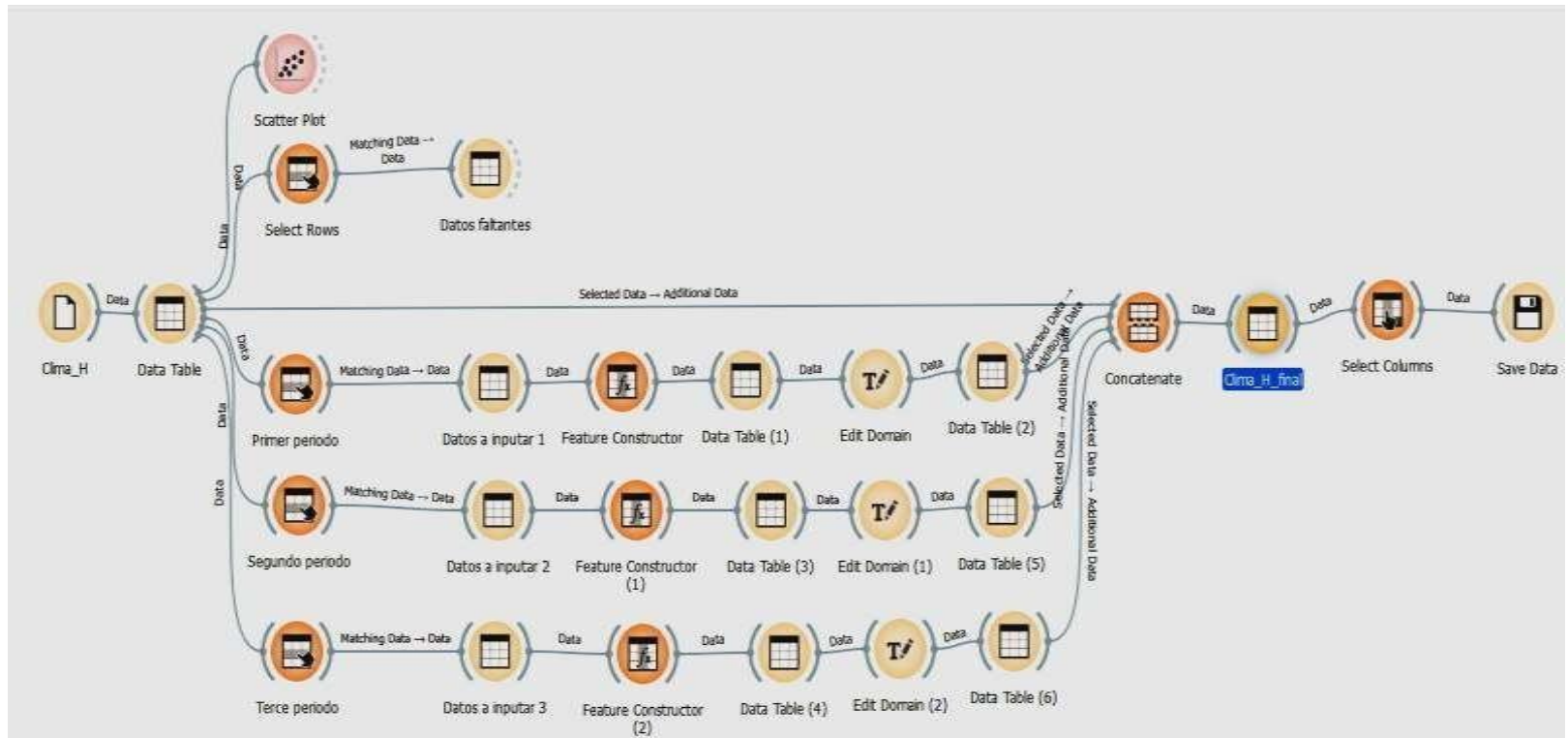


Figura 31 Modelo final de inputación de datos

La Fig. 31 muestra el uso del widget "**Concatenate**" para agregar los datos faltantes al conjunto de datos inicial *Clima_H cargado*. Este modelo combina de manera efectiva los datos de los periodos faltantes con el conjunto de datos original, resultando en el conjunto de datos final denominado *Clima_H_final.csv*. La imputación de datos y la consolidación del conjunto de datos son pasos cruciales en la preparación de datos antes de realizar análisis y predicciones.

El widget "*Feature Constructor*" se utiliza para realizar el cálculo de reducir el año en los registros de datos para que coincidan con los periodos que se van a completar. Esto es importante para garantizar que los datos sean consistentes y coincidan en el tiempo con los periodos de datos faltantes que se están imputando. Este proceso contribuye a la preparación adecuada de los datos antes de su análisis y predicción.

	Fecha	Ptemp	Radiacion	Pvelviento	Pdirviento
1	2019-07-01 15:0...	23.864	394.009	0.734652	190.582
2	2019-07-01 16:0...	23.3301	189.997	1.73089	163.238
3	2019-07-01 17:0...	19.5332	36.0344	2.95715	166.871
4	2019-07-01 18:0...	16.471	0	2.12576	159.482
5	2019-07-01 19:0...	14.6669	0	2.04222	152.386
6	2019-07-01 20:0...	13.4795	0	1.9476	166.073
7	2019-07-01 21:0...	13.224	0	2.46924	164.905
8	2019-07-01 22:0...	13.105	0	2.00705	164.625
9	2019-07-01 23:0...	13.0419	0	1.54045	168.919
10	2019-07-02 00:0...	12.9156	0	1.2417	156.575

Figura 32 Datos concatenados

En la Fig. 32, se puede comprobar que el archivo *Clima_H_final* contiene los datos completos. Se realiza una comprobación visualizando los registros de cada conjunto de datos. En el conjunto de datos *Clima_H* de la Figura 23, se tienen 18,304 registros, mientras que en el conjunto de datos *Clima_H_final*, con los datos imputados, se cuentan 24,105 registros. Esto demuestra que se han completado con éxito los datos faltantes en el conjunto de datos *Clima_H_final*.

Es importante confirmar la integridad del conjunto de datos mediante una gráfica de puntos para verificar que todos los datos estén completos. Comparando con la Fig. 25, se puede comprobar en esta gráfica de la Fig. 33 los resultados finales. Esto asegura que los datos faltantes han sido exitosamente imputados y que el conjunto de datos está completo y listo para su análisis y modelado.

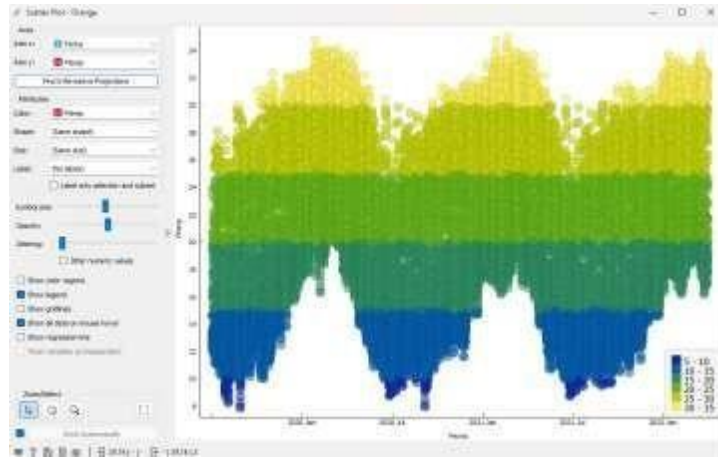


Figura 33 Visualización gráfica de puntos para la temperatura

En la Fig. 34, se comprueba efectivamente que la data esta completa para ello se empleó el widget Scatter Plot o gráfica de puntos. Similarmente se puede seleccionar las demás variables en la opción *Axis y*. Con lo cual se finaliza la fase de preparación de los datos.

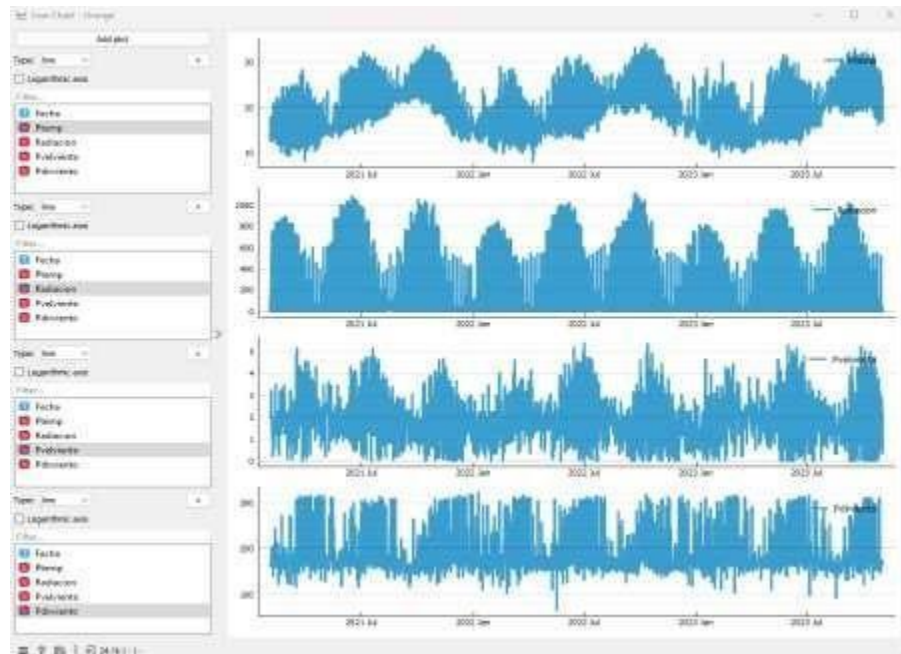


Figura 34 Comportamiento de las variables meteorológicas

El análisis de las variables meteorológicas en estudio, a través de gráficos de series de tiempo con representación lineal, revela un patrón estacional con una periodicidad evidente.

Por lo tanto: La Figura 34 confirma la integridad de los datos, demostrando que la información está completa, utilizando el widget Scatter Plot o gráfica de puntos para visualizar la totalidad de los registros.

Esta fase de preparación de datos ha sido esencial para asegurar la validez y confiabilidad de los análisis subsiguientes. El uso del widget Scatter Plot ha permitido verificar la integridad de la data y garantizar que no haya huecos o datos faltantes que puedan afectar la calidad de los resultados. Además, la capacidad de seleccionar otras variables en la opción Axis proporciona flexibilidad para extender este análisis a otras características meteorológicas.

Los gráficos de series de tiempo con representación lineal han proporcionado una visión clara de los patrones estacionales, revelando una periodicidad distintiva en las variables meteorológicas estudiadas. Este análisis inicial establece una base sólida para la aplicación de modelos de inteligencia artificial y la evaluación de su impacto en la predicción y comprensión del clima en la región de estudio durante el período considerado, como se puede apreciar en la visualización gráfica de puntos para la temperatura.

3.3. Predecir los parámetros meteorológicos y la inteligencia artificial en los fenómenos atmosféricos en el “Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022”.

Para la predicción de las variables meteorológicas, mediante la inteligencia artificial se utilizaron los algoritmos de series de tiempo *ARIMA Model* y *VAR Model*. Estos algoritmos se aplicaron por separado a cada una de las variables meteorológicas.

Antes de aplicar los modelos de predicción a la variable de temperatura, se llevó a cabo una evaluación de los modelos utilizando el widget *Model Evaluation*. Luego, se procedió a realizar la evaluación para la variable de temperatura.

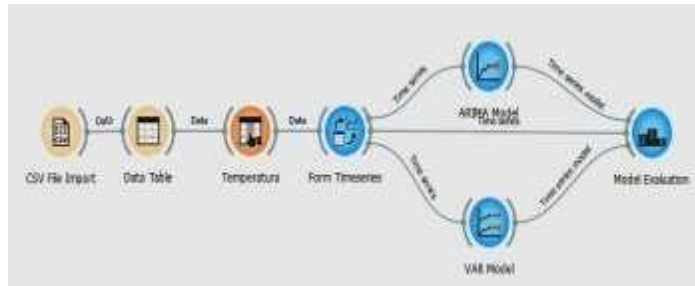


Figura 35 Modelo de evaluación para los algoritmos de series de tiempo

La Figura 35 muestra el modelo de evaluación de los dos algoritmos utilizados para la predicción de las variables meteorológicas, que son *ARIMA Model* y *VAR Model*.

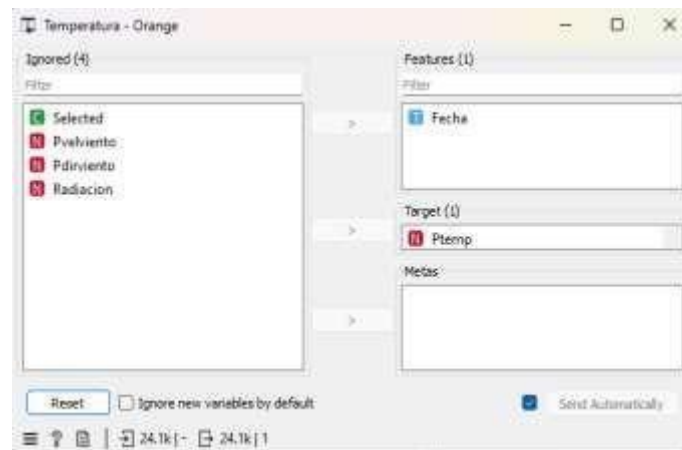


Figura 36 Selección de variables de predicción

La Fig. 36 muestra el uso del *Widget Data Table* para seleccionar la variable que se va a evaluar en términos de predicción, en este caso, la variable "*Temperatura*" (Target).

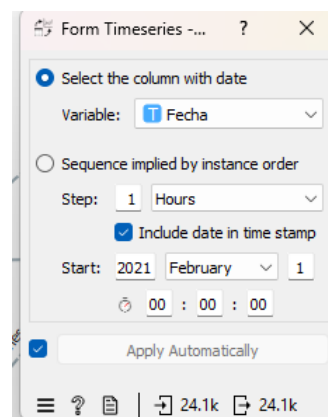
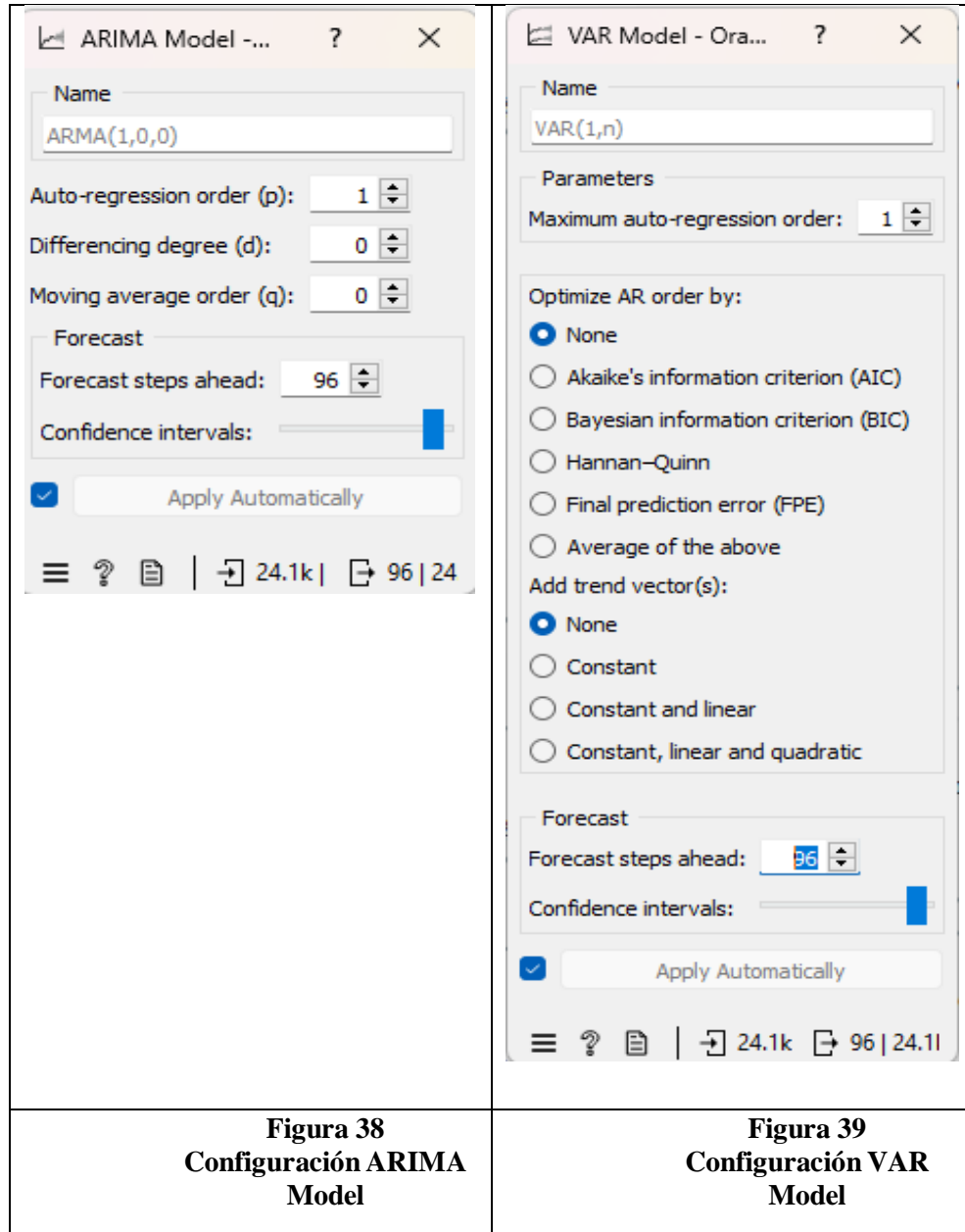


Figura 37 Configuración serie de tiempo

El Widget para series de tiempo, llamado "*Form Timeseries*", debe configurarse con la variable de tiempo y la secuencia de pasos en el tiempo. En su caso, la secuencia de tiempo se expresó en horas para su análisis.



En las Figs. 38 y 39, se puede observar la configuración de los modelos para que la predicción se realice en un horizonte de tiempo de 96 pasos, lo que equivale a 4 días, con un período de 24 horas.

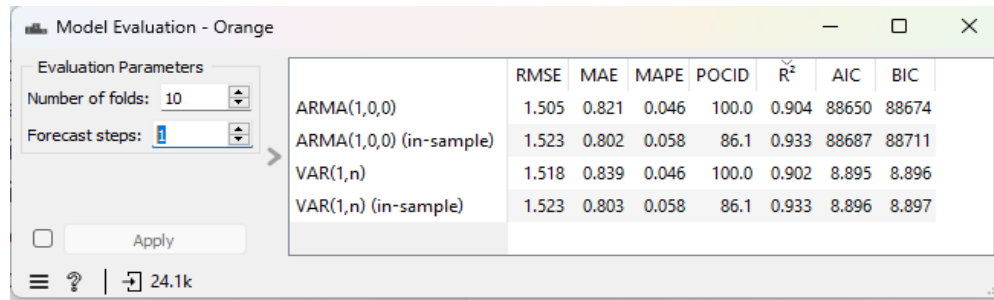


Figura 40 “Resultados de la evaluación de los modelos” ARIMA Model y VAR Model

En la Fig. 40, se muestran la evaluación de los modelos es necesaria para determinar su eficacia ARIMA y VAR, que incluye la comparación de los errores en términos de error cuadrático medio (RMSE), error absoluto medio (MAE), error porcentual absoluto medio (MAPE), predicción del cambio de dirección (POCID), coeficiente de determinación (R^2), criterio de información de Akaike (AIC) y criterio de información bayesiano (BIC).

Según el coeficiente de determinación (R^2), el modelo ARIMA obtuvo un valor de 90.4%, y el modelo VAR obtuvo un valor de 90.2%. Estos valores indican la capacidad de los modelos para explicar la variabilidad en los datos observados. Ambos modelos tienen un alto nivel de ajuste a los datos, lo que sugiere que son adecuados para la predicción de las variables meteorológicas

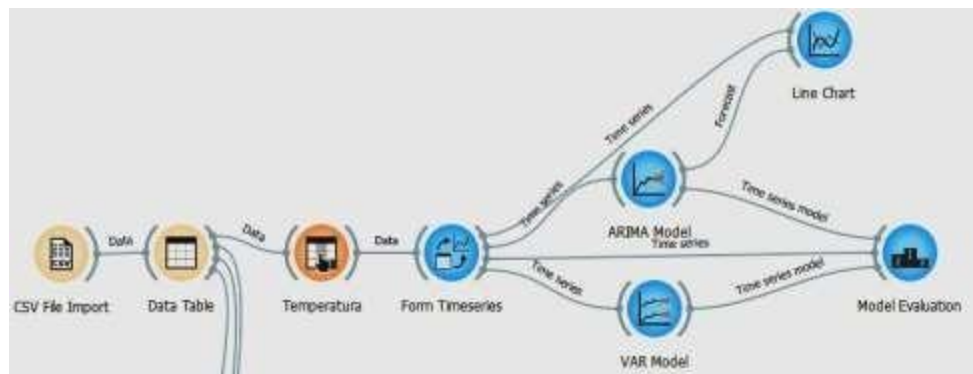


Figura 41 Predicción del modelo ARIMA Model

El Widget Line Chart se utiliza para mostrar la predicción realizada por el modelo **ARIMA Model**. Puede proporcionar una representación visual de cómo se compara la predicción con los datos observados y evaluar la precisión del modelo.

El modelo **ARIMA Model** es eficiente para la predicción de la temperatura, pero se debe aplicar a las demás variables meteorológicas, como la velocidad del viento, la radiación solar, la humedad relativa, etc. Luego, se puede utilizar el **widget Line Chart** para visualizar las predicciones de estas variables y compararlas con los datos observados. Esto permitirá evaluar la capacidad del modelo ARIMA para predecir diferentes aspectos del clima en la región de Ica.

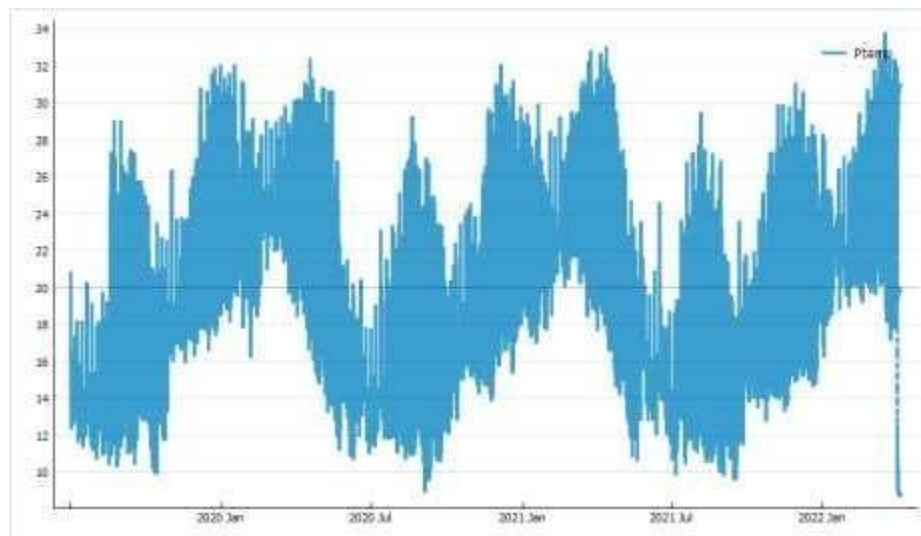


Figura 42 Predicción de las temperaturas

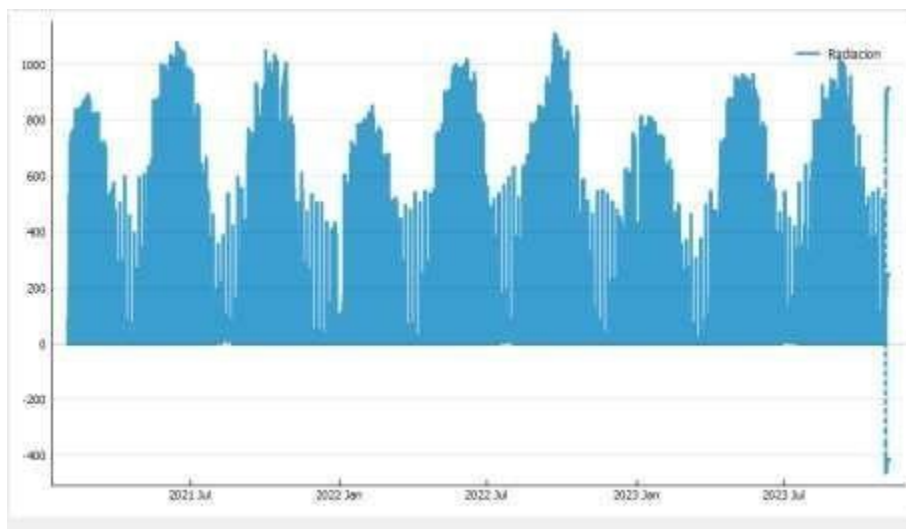


Figura 43 Predicción de la Radiación

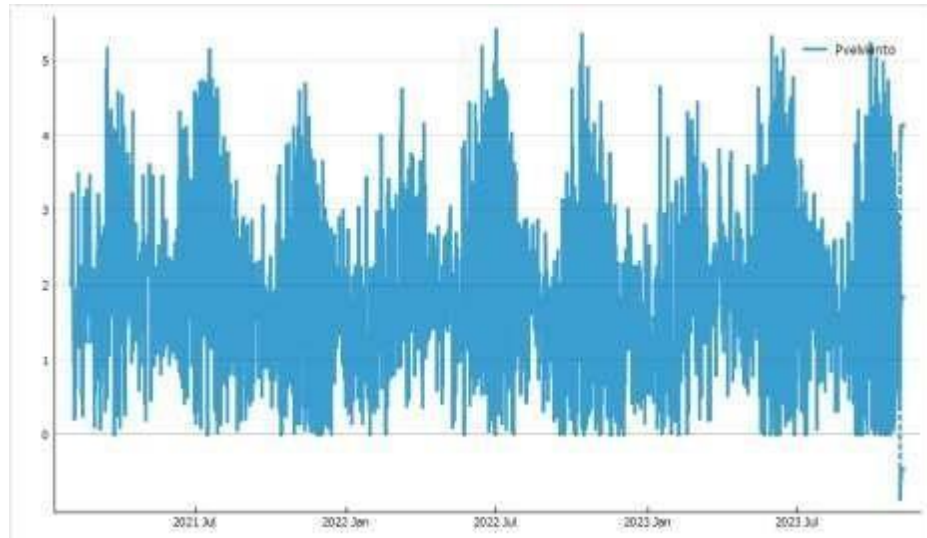


Figura 44 Predicción de la velocidad del viento

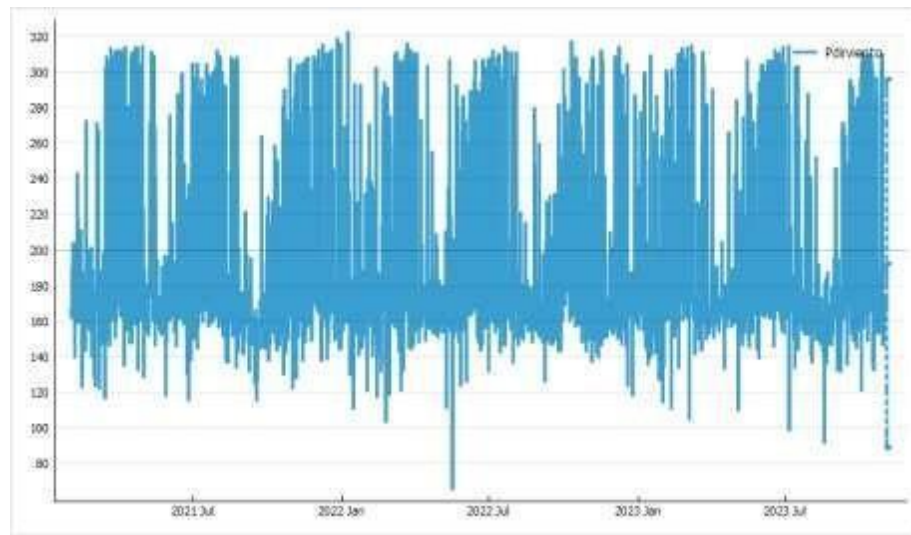


Figura 45 Predicción de la dirección del viento

Por lo tanto: sea abordado de manera exitosa. La eficiencia demostrada por el **modelo ARIMA** en la predicción de la temperatura establece una base sólida para extender este enfoque a otras variables meteorológicas cruciales, como la velocidad del viento, la radiación solar, la humedad relativa, entre otras.

La aplicación del **modelo ARIMA** a estas variables adicionales, seguida por la visualización de las predicciones a través del **widget Line Chart** y la comparación con los datos observados, permite una evaluación integral de la capacidad del modelo para

prever diversos aspectos del clima en la región de Ica. Este análisis permitirá determinar en qué medida la inteligencia artificial contribuye a la anticipación y prevención efectiva de fenómenos atmosféricos, proporcionando información valiosa para la gestión y mitigación de riesgos climáticos en la zona de estudio. La combinación de modelos precisos y visualizaciones claras fortalecerá la comprensión y aplicación de la inteligencia artificial en la gestión climática de la región.

En las pruebas de evaluación de los modelos *ARIMA Model* y *VAR Model*, se tienen como resultados para los diversos parámetros de evaluación y que se resumen en la Tabla X, concordantes con la Fig. 38.

TABLA 2 RESULTADOS DE LA EVALUACIÓN DE LOS MODELOS DE PREDICCIÓN

Modelo	RMSE	MAE	MAPE	POCID	R²	AIC	BIC
ARIMA (1,0,0)	1.505	0.821	0.046	100.0	0.904	88650	88674
VAR (1,n)	1.518	0.839	0.046	100.0	0.902	8.895	8.896

Según lo evaluado el modelo ARIMA Model supera al modelo VAR Model en términos de Error Cuadrático Medio (RMSE), Error Absoluto Medio (MAE) y Coeficiente de Determinación (R²). Estos resultados sugieren que el modelo ARIMA Model es más eficiente y preciso en la predicción de las variables meteorológicas en comparación con el modelo VAR Model.

IV. DISCUSIÓN

4.1. **Discusión de resultados: modelamiento de los parámetros meteorológicos y la inteligencia artificial en el clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022.**

La comprensión del problema y la exploración de datos son etapas cruciales en cualquier proyecto de modelamiento, especialmente en el ámbito de la meteorología mediante inteligencia artificial. La metodología CRISP-DM proporcionó una estructura sólida para abordar estas fases, permitiendo una comprensión detallada del problema y la identificación de huecos en los datos. Los resultados revelaron, a través de las gráficas generadas con los widgets "From Timeseries" y "Line Chart", la presencia de espacios vacíos o huecos en los datos, señalados con flechas de color rojo en la Figura 13.

Estos vacíos en los datos se consideraron elementos críticos que requerían corrección antes de aplicar modelos de análisis y predicción. Como destacan *Mercado et al.*, la integridad de los datos es esencial para garantizar la calidad y precisión de los modelos, y la presencia de datos faltantes podría comprometer estos aspectos. Por lo tanto, la corrección de estos vacíos se volvió imperativa para asegurar la fiabilidad de los modelos utilizados en el estudio[28].

En este contexto, la implementación de técnicas de imputación de datos, como se describió en las secciones anteriores, desempeñó un papel crucial en la mejora de la integridad del conjunto de datos. Al abordar y corregir estos vacíos, se fortaleció la base sobre la cual se realizaron los análisis y predicciones, mejorando así la confiabilidad de los resultados obtenidos. Este enfoque se alinea con las mejores prácticas en el manejo de datos meteorológicos para garantizar la validez de los modelos y sus aplicaciones en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra en Ica, durante el período 2019-2022.

4.2. Discusion de resultados: elemento temperatura y la inteligencia artificial en el clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022.

El proceso de preparación de datos, especialmente en relación con las variables de temperatura y radiación, desempeña un papel crucial en la calidad y robustez de los modelos de inteligencia artificial aplicados al clima en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra en Ica. La transformación de unidades a una escala temporal uniforme, como se ilustra en la Fig. 16, demuestra un enfoque meticuloso para estandarizar las variables, facilitando así su comparación y análisis coherente.

La estrategia de imputación de datos empleada, como se describe en la Fig. 25, donde se han identificado y abordado específicamente los periodos con datos faltantes, demuestra un enfoque riguroso para preservar la integridad temporal de la serie. La detección de los intervalos específicos con datos faltantes, marcados en la Fig. 26, proporciona una base sólida para la imputación precisa y contribuye a la coherencia de la serie temporal completa.

Este nivel de detalle en la preparación de datos es esencial para garantizar la calidad de los análisis y predicciones subsiguientes. La detección y corrección de datos faltantes no solo mejora la integridad del conjunto de datos, sino que también respalda la robustez de los modelos de inteligencia artificial al reducir la posibilidad de sesgos o distorsiones en las proyecciones climáticas. Estos procedimientos metodológicos sólidos son cruciales para mantener la fiabilidad de los resultados y contribuir significativamente al entendimiento y la predicción del clima en la región de Ica.

4.3. Discusión de resultados: los parámetros meteorológicos y la inteligencia artificial en los fenómenos atmosféricos en el Centro de Investigación del Estudio de la Actividad Solar y sus Efectos Sobre la Tierra, Ica, 2019-2022.

La discusión de los resultados obtenidos de la aplicación de los modelos ARIMA y VAR Model proporciona insights valiosos sobre la eficacia de la inteligencia artificial en la predicción de variables meteorológicas. Los resultados presentados en la Tabla X reflejan métricas clave de

desempeño, y la comparación de estos indicadores revela que el modelo ARIMA (1,0,0) supera al VAR (1,n) en términos de precisión predictiva.

El valor del Error Cuadrático Medio (RMSE) y del Error Absoluto Medio (MAE) más bajo para el modelo ARIMA sugiere que este modelo tiene una capacidad superior para minimizar la discrepancia entre las predicciones y los valores observados. Además, el Coeficiente de Determinación (R^2) cercano al 0.904 para ARIMA indica un alto grado de ajuste del modelo a los datos reales, destacando su capacidad para explicar la variabilidad en las series temporales.

El análisis de los resultados también muestra que el modelo VAR presenta métricas de desempeño competitivas, pero ligeramente inferiores en comparación con ARIMA. Esto puede atribuirse a la naturaleza específica de las variables meteorológicas estudiadas y cómo interactúan en el entorno.

V. CONCLUSIONES

1. Se concluye que los resultados obtenidos revelaron la presencia de huecos o espacios vacíos en los datos, destacados mediante gráficas generadas con widgets específicos, como "From Timeseries" y "Line Chart".

Estos vacíos en los datos fueron identificados como elementos críticos que requerían corrección antes de aplicar modelos de análisis y predicción. La discusión se enfocó en la importancia de abordar y corregir estos datos faltantes para garantizarla.

2. Se concluye, que durante el período 2019-2022, es fundamental para comprender y prever el comportamiento climático. La metodología CRISP-DM proporcionó una estructura robusta para abordar la complejidad de los datos climatológicos y la comprensión del problema, destacando la importancia de la fase de preparación de datos.

La detección y corrección de datos faltantes, especialmente en variables críticas como la temperatura y la radiación, son pasos esenciales para preservar la integridad de la serie temporal y garantizar la coherencia en los análisis. La imputación estratégica de datos perdidos contribuye significativamente a la completitud del conjunto de datos y mejora la calidad de los resultados.

En este contexto, la implementación del modelo ARIMA ha demostrado ser eficiente para predecir la temperatura.

3. En conclusión, los resultados de la aplicación de los modelos ARIMA y VAR Model en la predicción de variables meteorológicas, indican que el modelo ARIMA (1,0,0) demostró ser más efectivo en términos de precisión y ajuste a los datos en comparación con el modelo VAR (1,n). La elección del ARIMA se respalda en métricas como el Error Cuadrático Medio (RMSE) y el Coeficiente de Determinación (R^2), que indican un rendimiento superior en la capacidad predictiva y la capacidad del modelo para explicar la variabilidad en las series temporales.

Estos hallazgos subrayan la relevancia de seleccionar cuidadosamente el modelo apropiado para abordar las complejidades y características específicas de los datos meteorológicos. La inteligencia artificial, en este caso representada por el modelo ARIMA, emerge como una

herramienta valiosa para prever fenómenos atmosféricos y contribuir significativamente al análisis del clima en la región de Ica. Sin embargo, se recomienda continuar explorando y refinando enfoques, considerando la posibilidad de modelos alternativos y ajustes específicos para las variables climáticas estudiadas.

Este análisis respalda la viabilidad de emplear la inteligencia artificial en la predicción climática, proporcionando perspectivas que pueden mejorar la capacidad de anticipación y gestión de fenómenos atmosféricos.

VI. RECOMENDACIONES

1. La recomendación se buscó centrar en fortalecer la calidad de los modelos de inteligencia artificial aplicados al modelamiento de parámetros meteorológicos. Se sugiere implementar procesos de validación y corrección periódicos de los datos, enfocándose especialmente en la identificación y tratamiento de huecos o datos faltantes. Además, se recomienda la exploración de técnicas avanzadas de inteligencia artificial, como redes neuronales, para mejorar la capacidad predictiva de los modelos. Se aconseja también mantener una vigilancia constante sobre la evolución de las condiciones climáticas y la aparición de patrones no previstos. La actualización regular de los modelos con datos en tiempo real contribuirá a su adaptabilidad y precisión.
2. En base a los resultados obtenidos, se formulan diversas recomendaciones para optimizar el uso de inteligencia artificial en el modelamiento de parámetros meteorológicos. En primer lugar, se sugiere extender la aplicación del modelo ARIMA a otras variables climáticas, como velocidad del viento, humedad relativa, y radiación solar, para obtener un panorama integral y detallado de las condiciones atmosféricas.

Asimismo, se hace hincapié en la importancia de continuar monitoreando y mejorando la calidad de los datos. La detección temprana de posibles anomalías en la recolección de datos y la implementación de protocolos para su corrección son fundamentales para garantizar la confiabilidad de los resultados y prevenir sesgos en el análisis climatológico.

También. Se recomienda explorar la integración de otros modelos de inteligencia artificial complementarios al ARIMA, como redes neuronales o modelos de aprendizaje profundo, para evaluar su eficacia en la predicción de variables meteorológicas específicas. La combinación de diferentes enfoques puede proporcionar una visión más completa y precisa de los patrones climáticos en la región.
3. Con base en los resultados discutidos, se derivan varias recomendaciones para futuras investigaciones y aplicaciones prácticas. En primer lugar, se sugiere explorar la incorporación de modelos más avanzados de inteligencia artificial, además del

ARIMA, que puedan capturar de manera más precisa las complejidades de las variables meteorológicas en la región de Ica. La inclusión de enfoques como redes neuronales o modelos de aprendizaje profundo podría ofrecer una comprensión más profunda de las relaciones subyacentes en los datos climáticos.

Además, se recomienda realizar un análisis más detallado de los períodos con datos faltantes, identificados como críticos en el proceso de imputación. Investigaciones adicionales podrían enfocarse en estrategias específicas para abordar y corregir estos vacíos de manera más efectiva, considerando factores estacionales o patrones climáticos particulares que puedan influir en la imputación de datos.

En términos de aplicación práctica, se insta a considerar la integración de los modelos de predicción desarrollados en sistemas de alerta temprana para fenómenos atmosféricos en la región. La implementación de tales sistemas podría mejorar la capacidad de respuesta y la toma de decisiones ante eventos climáticos adversos, contribuyendo así a la gestión de riesgos y la protección de la infraestructura y la población local.

VII. REFERENCIAS BIBLIOGRÁFICAS

- [1] J. C. Lennox 2084, «Artificial Intelligence and the Future of Humanity», *Rev. Fe y Lib.*, vol. 4, n.º 1, pp. 207-209, 2021, [En línea]. Disponible en: <file:///C:/Users/Usuario/Downloads/109-Texto del artículo-363-1-10-20211215.pdf>.
- [2] M. Garrido Peña, «Machine Learning y Data Science con Python», *Udemy*, 2020. <https://www.udemy.com/course/aprende-data-science-y-machine-learning-con-python/learn/lecture/7789510#questions>.
- [3] Andrew Ng, *Coursera Curso: Machine Learning, Universidad de Stanford*. Universidad de Stanford., 2020.
- [4] R. Castro, Y. M. Souto, E. Ogasawara, F. Porto, y E. Bezerra, «STConvS2S: Spatiotemporal Convolutional Sequence to Sequence Network for weather forecasting», *Neurocomputing*, vol. 426, pp. 285-298, 2020, doi: 10.1016/j.neucom.2020.09.060.
- [5] «Building Tangent-Linear and Adjoint Models for Data Assimilation With Neural Networks», *J. Adv. Model. Earth Syst.*, vol. 13, n.º 9, pp. 1-16, 2021, doi: 10.1029/2021MS002521.
- [6] P. Hewage, M. Trovati, E. Pereira, y A. Behera, «Deep learning-based effective fine-grained weather forecasting model», *Pattern Anal. Appl.*, vol. 24, n.º 1, pp. 343-366, 2021, doi: 10.1007/s10044-020-00898-1.
- [7] S. Kim, J. Lee, y T. Yoon, «Road surface conditions forecasting in rainy weather using artificial neural networks», *Saf. Sci.*, vol. 140, n.º April, p. 6 Pag., 2021, doi: 10.1016/j.ssci.2021.105302.
- [8] Plex Scholar is a BETA, «LSTM Neural Network - Springer Nature». 2021.
- [9] X. Ren *et al.*, «Deep Learning-Based Weather Prediction: A Survey», *Big Data Res.*, vol. 23, p. 100178, 2021, doi: 10.1016/j.bdr.2020.100178.
- [10] H. J. Song y S. Roh, «Improved Weather Forecasting Using Neural Network Emulation for Radiation Parameterization», *J. Adv. Model. Earth Syst.*, vol. 13, n.º 10,

- pp. 1-21, 2021, doi: 10.1029/2021MS002609.
- [11] M. Nsabagwa, M. Byamukama, E. Kondela, y J. S. Otim, «Towards a robust and affordable Automatic Weather Station», *Dev. Eng.*, vol. 4, n.º May 2018, p. 100040, 2019, doi: 10.1016/j.deveng.2018.100040.
- [12] T. Dube, P. Moyo, M. Ncube, y D. Nyathi, «The Impact of Climate Change on Agro-Ecological Based Livelihoods in Africa: A Review», *J. Sustain. Dev.*, vol. 9, n.º 1, p. 256, 2016, doi: 10.5539/jsd.v9n1p256.
- [13] D. N. Mubiru, F. B. Kyazze, M. Radeny, A. Zziwa, J. Lwasa, y J. Kinyangi, *Climatic trends, risk perceptions and coping strategies of smallholder farmers in rural Uganda*. Uganda: Working Paper No. 121 CGIAR, 2015.
- [14] S. Lwasa, F. Mugagga, B. Wahab, D. Simon, J. Connors, y C. Griffith, «Urban and peri-urban agriculture and forestry: Transcending poverty alleviation to climate change mitigation and adaptation», *Urban Clim.*, vol. 7, pp. 92-106, 2014, doi: 10.1016/j.uclim.2013.10.007.
- [15] D. Kreuzer, M. Munz, y S. Schlüter, «Short-term temperature forecasts using a convolutional neural network — An application to different weather stations in Germany», *Mach. Learn. with Appl.*, vol. 2, n.º June, p. 100007, 2020, doi: 10.1016/j.mlwa.2020.100007.
- [16] R. Cavalier *et al.*, «Conditions for a market uptake of climate services for adaptation in France», *Clim. Serv.*, vol. 6, pp. 34-40, 2017, doi: 10.1016/j.cliser.2017.06.010.
- [17] R. L. Vilca, «Redes neuronales y lógica difusa aplicado a un sistema climatológico», Universidad Nacional San Agustín, 2017.
- [18] C. E. Armas Morales, «La Inteligencia Artificial en empresas peruanas e impactos laborales en los trabajadores», *Iberoam. Bus. J.*, vol. 5, n.º 1, pp. 83-105, 2021, doi: 10.22451/5817.ibj2021.vol5.1.11053.
- [19] G. Uscumayta Huaman, «Inteligencia Artificial como Herramienta Para Mejorar la Calidad y Celeridad de DEcisiones Jurisdiccionales», Universidad Andina del Cusco, 2021.

- [20] S. H. Bari, M. T. Rahman, M. M. Hussain, y S. Ray, «Forecasting Monthly Precipitation in Sylhet City Using ARIMA Model», *Civ. Environ. Res.*, vol. 7, n.º 1, pp. 69-77, 2015, [En línea]. Disponible en: <https://iiste.org/Journals/index.php/CER/article/view/19069>.
- [21] J. Castillo, «Organización Meteorología Mundial.», *Estaciones Meteorológicas*, 2017. <https://public.wmo.int/es>.
- [22] O. Peña, «Sistema De Información De Mini Estación Metereológica Para La Captura De Datos De Temperatura, Humedad Y Presión Atmosférica - Centro De Operaciones De Emergencia Regional - Piura», Universidad Nacional de Piura, 2019.
- [23] F. Zapata, «¿Qué es una estación meteorológica?», *Inst. PCE*, 2016. <https://www.pce-iberica.es/medidor-detalles-tecnicos/que-estacion-meteorologica.htm>.
- [24] C. Campetella, B. Cerne, y P. Salio, «Entornos invisibles, estación metereológica», en *Ciencia y Tecnología*, Argentina: Ministerio de Educación - Instituto Nacional de Educación Tecnológica, 2011, pp. 1-52.
- [25] S. P. Durrani, S. Balluff, L. Wurzer, y S. Krauter, «Photovoltaic yield prediction using an irradiance forecast model based on multiple neural networks», *J. Mod. Power Syst. Clean Energy*, vol. 6, n.º 2, pp. 255-267, 2018, doi: 10.1007/s40565-018-0393-5.
- [26] H. Kanchev, L. D., F. Colas, V. Lazarov, y B. Francois, «Energy management and operational planning of a microgrid with a PV-based active generator for smart grid applications», *IEEE Trans. Ind. Electron.*, vol. 58, n.º 10, pp. 4583-4592, 2011, doi: 10.1109/TIE.2011.2119451.
- [27] R. Sabzehgar, D. Z. Amirhosseini, y M. Rasouli, «Solar power forecast for a residential smart microgrid based on numerical weather predictions using artificial intelligence methods», *J. Build. Eng.*, vol. 32, n.º June, p. 101629, 2020, doi: 10.1016/j.job.2020.101629.
- [28] F. R. Mercado, W. García Fernández, y J. A. H. Acebey, «Sistema de inteligencia artificial para la predicción temprana de heladas meteorológicas Artificial intelligence system for early prediction of weather frost», *Acta Nov.*, vol. 7, n.º December, p. 13 Pag., 2016, [En línea]. Disponible en: file:///C:/Users/Usuario/Downloads/full_text

(2).pdf.

- [29] B. Mohammadi *et al.*, «Application of ERA-Interim, empirical models, and an artificial intelligence-based model for estimating daily solar radiation», *Ain Shams Eng. J.*, vol. 13, n.º 1, p. 12 Pag., 2021, doi: 10.1016/j.asej.2021.05.012.
- [30] İ. Uçkan y K. M. Khudhur, «Estimation and comparison of various global solar radiation models on horizontal surface in a hot and dry region», *Int. J. Green Energy*, vol. 15, n.º 5, pp. 358-370, 2018, doi: 10.1080/15435075.2018.1454322.
- [31] J. Zhang, L. Zhao, S. Deng, W. Xu, y Y. Zhang, «A critical review of the models used to estimate solar radiation», *Renew. Sustain. Energy Rev.*, vol. 70, n.º July 2016, pp. 314-329, 2017, doi: 10.1016/j.rser.2016.11.124.
- [32] O. A. Olanike y A. M. Chandra, «Evaluation of solar radiation estimation methods for reference evapotranspiration estimation in Canada», *Theor. Appl. Climatol. Vol.*, pp. 377–385, 2014, doi: <https://doi.org/10.1007/s00704-013-1070-2>.
- [33] B. . Jahani y B. Mohammadi, «A comparison between the application of empirical and ANN methods for estimation of daily global solar radiation in Iran», *Theor. Appl. Climatol.*, vol. 137, n.º 31, 2019, doi: 10.1007/s00704-018-2666-3.
- [34] Y. Liu, Y. Zhou, D. Wang, Y. Wang, Y. Li, y Y. Zhu, «Classification of solar radiation zones and general models for estimating the daily global solar radiation on horizontal surfaces in China», *Energy Convers. Manag.*, vol. 154, n.º November, pp. 168-179, 2017, doi: 10.1016/j.enconman.2017.10.043.
- [35] K. Yang, T. Koike, y B. Ye, «Improving estimation of hourly, daily, and monthly solar radiation by importing global data sets», *Agric. For. Meteorol.*, vol. 137, n.º 1-2, pp. 43-55, 2006, doi: 10.1016/j.agrformet.2006.02.001.
- [36] L. Wang *et al.*, «Prediction of diffuse solar radiation based on multiple variables in China», *Renew. Sustain. Energy Rev.*, vol. 103, n.º December 2018, pp. 151-216, 2019, doi: 10.1016/j.rser.2018.12.029.
- [37] X. Liu, X. Mei, Y. Li, J. R. Porter, Q. Wang, y Y. Zhang, «Choice of the Ångström-Prescott coefficients: Are time-dependent ones better than fixed ones in modeling global solar irradiance?», *Energy Convers. Manag.*, vol. 51, n.º 12, pp. 2565-2574,

- 2010, doi: 10.1016/j.enconman.2010.05.020.
- [38] B. Mohammadi y S. Mehdizadeh, «Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm», *Agric. Water Manag.*, vol. 237, n.º March, p. 106145, 2020, doi: 10.1016/j.agwat.2020.106145.
- [39] B. Yaniktepe, O. Kara, y C. Ozalp, «The global solar radiation estimation and analysis of solar energy: Case study for Osmaniye, Turkey», *Int. J. Green Energy*, vol. 14, n.º 9, pp. 765-773, 2017, doi: 10.1080/15435075.2017.1329148.
- [40] V. Laqui, «Efecto de la variación de los elementos climaticos en la producción de quinua, en la cuenca del rio Ilave-Puno», Universidad Nacional del Altiplano, 2017.
- [41] J. Hansen, M. Sato, y R. Ruedy, «Perception of climate change», *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, n.º 37, 2012, doi: 10.1073/pnas.1205276109.
- [42] INEI, *Instituto Nacional de estadística e Informatica. Sistema ESTADISTICO nacional*. Oficina Departamental de Estadística e Informatica de ICA, 2017.
- [43] R. Hernandez, C. Fernandez, y P. Baptista, *Metodologia de la Investigacion*, Sexta Edic. Mexico: Miembro de la Cámara Nacional de la Industria Editorial Mexicana, Reg. Núm. 736, 2014.
- [44] C. Weng, R. B. Ghazali, S. A. Mustafa, A. N. Kareem, y B. A. Khalaf, «Weather Forecasting for Batu Pahat Using Neural Network», *Fusion Pract. Appl.*, n.º January, pp. 64-70, 2021, doi: 10.54216/fpa.060204.