




PRONÓSTICO DE PARÁMETROS METEOROLÓGICOS, USANDO MINERÍA DE DATOS. UN CASO PRÁCTICO

FORECAST OF METEOROLOGICAL PARAMETERS, USING DATA MINING. A PRACTICAL CASE

 Jesús Espinola Gonzales¹  Ángel Cobo Ortega²  Rocío Rocha Blanco²

¹Universidad Tecnológica de Los Andes, Perú.

²Universidad de Cantabria, España.

Correspondencia:

Dr. Jesús Espinola Gonzales

jespinolag@utea.edu.pe

Para citar este artículo: Espinola, J., Cobo, Á., & Rocha, R. (2022). Pronóstico de parámetros meteorológicos, usando minería de datos. Un caso práctico. *Revista de Investigación Hatun Yachay Wasi*, 1(1), 112 - 127. <https://doi.org/10.57107/hyw.v1i1.15>

RESUMEN

Este trabajo aborda el problema del pronóstico de parámetros meteorológicos como son temperatura, humedad relativa, radiación solar, punto de rocío y precipitaciones, a partir de un conjunto de datos obtenidos en una estación meteorológica. Para resolver este problema se han usado técnicas de minería de datos, entre ellas M5Rules, M5P, Regresión lineal y Vecino más cercano. Como resultado del trabajo se muestra que la técnica M5Rules es apropiada para el pronóstico de temperatura, humedad relativa, así como para punto de rocío; mientras que la técnica del vecino más cercano es más apropiada para el pronóstico de radiación solar y precipitaciones.

Palabras clave: Pronóstico, parámetros meteorológicos, minería de datos.

ABSTRACT

This work addresses the problem of forecasting meteorological parameters such as temperature, relative humidity, solar radiation, dew point and rainfall, from a set of data obtained from a weather station. In order to solve this problem, data mining techniques have been used, including M5Rules, M5P, Linear Regression and Nearest Neighbor. As a result, it is shown that the M5Rules technique is appropriate for the forecast of temperature, relative humidity, as well as for dew point; while the nearest neighbor technique is more appropriate for forecasting solar radiation and precipitation.

Keywords: Forecast, meteorological parameters, data mining.



INTRODUCCIÓN

Tener el pronóstico (tendencias) a partir de un conjunto de datos es una necesidad en muchas áreas del saber, por lo que es un problema frecuente a resolver. Particularmente también lo es el pronóstico de parámetros meteorológicos: temperatura, humedad relativa, radiación solar, punto de rocío y precipitaciones; pues tiene particular importancia para el turismo, la agricultura, y muchas otras actividades económicas y sociales. Los parámetros meteorológicos son los elementos que tienen como finalidad caracterizar el estado del tiempo, estos son medibles y logran determinar las condiciones presentes en la atmósfera (Eraso, 2020).

En el Perú se tiene información del pronóstico del tiempo que proporciona el Servicio Nacional de Meteorología e Hidrología del Perú [SENAMHI] (SENAMHI, s.f.), pero esta información no está particularizada para lugares geográficos puntuales, sino para áreas geográficas más amplias y en ocasiones difícilmente extrapolable a áreas con especiales características orográficas o ambientales.

La información de los parámetros meteorológicos que se recoge en las estaciones meteorológicas es abundante, por lo que para analizarla adecuadamente se requieren técnicas especiales; entre ellas destacan las que ofrece la minería de datos, como por ejemplo la generación de reglas con M5Rules, M5P, la regresión lineal o las técnicas de clasificación y

predicción basadas en el concepto de vecino más cercano. En este contexto, la minería de datos es una herramienta estratégica clave para explotar el conocimiento presente en los datos, permite analizarlos desde todas las perspectivas, con el fin de transformar la información (Haro et al., 2018).

Entre la amplia variedad de herramientas informáticas que implementan técnicas de tratamiento y minería de datos, en este trabajo se ha hecho uso de Pentaho (Stratebi, s.f.), una importante aplicación en el área de inteligencia de negocios. También se ha usado la aplicación Weka, muy conocida y usada en el campo de la minería de datos.

Martínez - Abad & Hernández - Ramos (2018), mostraron el potencial del programa estadístico Weka para análisis estadístico de información masiva a partir de bases de datos con grandes volúmenes de información.

MATERIALES Y MÉTODOS

El objetivo de este estudio es, utilizar técnicas de minería de datos para el pronóstico de parámetros meteorológicos. Como caso práctico se aborda el problema de determinar el pronóstico de los parámetros meteorológicos a partir de un conjunto de datos obtenidos en una estación meteorológica en Shancayán – Huaraz - Perú. El proceso de obtención de esos datos exigió la automatización de la captación, procesamiento y aplicación de técnicas analíticas.

Las fases para la recolección y procesamiento de datos es la que se establece en la metodología SEMMA, que es la de: muestreo, exploración, modificación, modelado y evaluación de los datos.

Recolección de datos

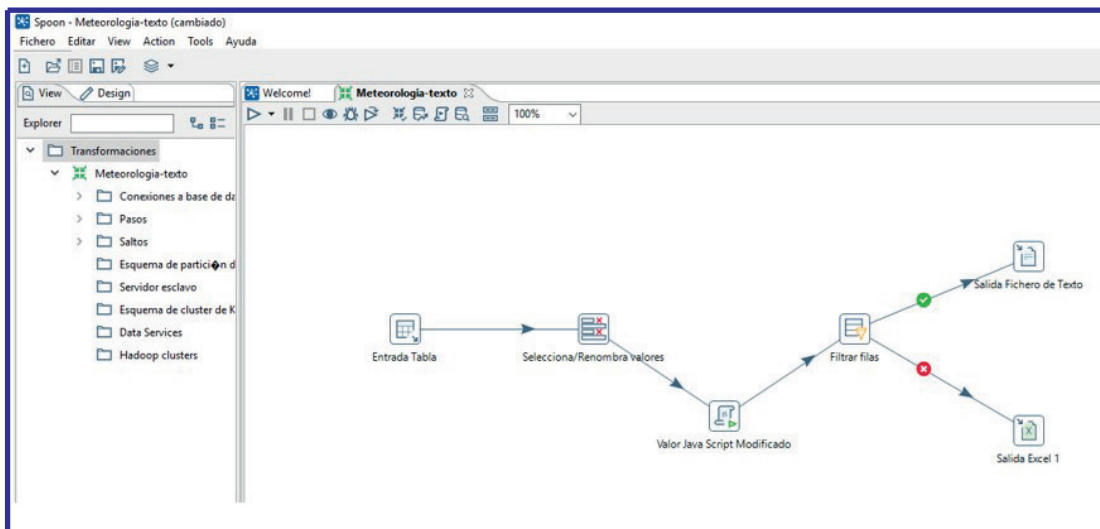
Los datos trabajados fueron proporcionados por el Centro de Investigación Ambiental para el Desarrollo de la Universidad Nacional “Santiago Antúnez de Mayolo”, datos que corresponden a la estación meteorológica en la zona de Shancayán – Huaraz, ubicada a 3050 m s.n.m. El estudio se ha hecho tomando como base 8730 registros, estos tienen la información, para cada hora en el periodo del 01 de diciembre de 2016 al 30 de noviembre de 2017, de los parámetros meteorológicos en estudio (temperatura, humedad relativa, radiación solar, punto de rocío y precipitación). Debe destacarse que se considera esta muestra de datos de un ámbito geográfico y temporal muy localizado, para valorar la potencia predictiva de las técnicas estudiadas,

potencia predictiva de las técnicas estudiadas, que posteriormente podría ser objeto de extrapolación a otros contextos similares.

Para acceder a la Base de Datos que almacena las mediciones de los sensores meteorológicos, se ha hecho uso de una herramienta de inteligencia de negocios, el software Pentaho, (Hitachi Vantara, 2022), y en concreto su módulo de integración de datos; el cual ha permitido extraer los datos registrados de los diferentes parámetros meteorológicos y ordenarlos cronológicamente, esto de una forma sencilla y en una sola fase. Este software, que permite integrar datos de diferentes fuentes, a su vez permite la salida de la información en diferentes formatos. En este trabajo se ha configurado que la salida sea en formato de texto para su posterior uso con la siguiente herramienta de procesamiento de los datos. La Figura 1, muestra el diseño del trabajo definido con la herramienta de extracción, transformación y carga de datos del paquete Pentaho.

FIGURA 1

Ventana de trabajo de Pentaho



Procesamiento de datos

El procesamiento de datos se ha realizado haciendo uso del software Weka, (Weka, s.f.). Esta aplicación tiene incorporadas diferentes técnicas de minería de datos. Entre estas técnicas destacan las predictivas, que permiten hacer pronósticos de variables de salida a partir de un conjunto de datos iniciales correspondientes a variables de entrada.

Exploración de los datos

Una vez obtenidos los datos desde la Base de Datos respectiva, se hizo una exploración de éstos que permitió ver ya ciertas relaciones entre algunos de los parámetros meteorológicos en estudio. Se vio por ejemplo que había una aparente relación directa entre la temperatura y la radiación solar y una aparente relación inversa entre la temperatura y la humedad relativa. También se vio la necesidad de subsanar algunos errores en los datos, debido a problemas en la transmisión de estos. De hecho, una de las fases que más tiempo y esfuerzo consume en un proyecto de minería de datos es el preprocesamiento de los datos, identificando posibles errores, datos ausentes, presencia de datos anómalos.

Modificación de los datos

En esta siguiente etapa se realizó la subsanación de los datos anómalos o ausentes, en este último caso mediante una interpolación entre los datos registrados en las horas previas y siguientes. También se completó el archivo de datos, en formato texto, para poder ponerlo en el formato de lectura de Weka.

Modelados

Para seleccionar las técnicas o algoritmos más apropiados para el pronóstico de los diferentes parámetros meteorológicos se realizaron pruebas con diferentes opciones que ofrece Weka, en concreto las siguientes:

1. M5Rules: técnica basada en árboles de decisión que utiliza estrategias iterativas para generar reglas que puedan aplicarse a problemas de clasificación y regresión.
2. M5P: algoritmo basado en árboles de regresión que puede ajustar automáticamente y eficientemente datos altamente no lineales y desestructurados.
3. Regresión lineal: método matemático que modela la relación entre una variable dependiente y una serie de variables independientes mediante una ecuación lineal.
4. Vecino más cercano (IBk): clasificador basado en instancias que utiliza la clase de las instancias de entrenamiento más próximas para deducir la clase de la instancia a clasificar.

De los análisis realizados en el contexto de este trabajo se deduce que para obtener mejores tendencias (pronósticos) de temperatura, humedad relativa y punto de rocío la técnica “M5-Rules” (Duggal & Singh, 2012), daba mejores resultados; sin embargo, para obtener mejores pronósticos de las precipitaciones y de la radiación solar el algoritmo del vecino más cercano (IBk) resultó el más apropiado.

Evaluación

Para evaluar cada una de las técnicas se utilizaron las series de datos correspondientes a las mediciones del periodo comprendido entre el 1 de diciembre del 2016 y el 30 de noviembre del 2017, para realizar un pronóstico del correspondiente parámetro meteorológico para las siguientes 24 horas (1 de diciembre de 2017). La calidad del pronóstico se analizó comparando los valores obtenidos con los medidos por la estación meteorológica y que se almacenaban en la base de datos original. La medición de la calidad se realizó considerando el error cuadrático medio. La siguiente sección resume los resultados obtenidos.

RESULTADOS Y DISCUSIÓN

Temperatura promedio del aire

La temperatura promedio registrada en la base de datos tenía un rango de variación amplio, entre 0.6 °C y 23.6 °C.

Utilizando las cuatro técnicas los resultados obtenidos son los que se muestran en la Tabla 1, donde al final se presenta el error cuadrático medio para cada técnica comparada con los datos reales. Se aprecia claramente que la técnica M5Rules y M5P para árboles dan mejores resultados. Esto confirma la propuesta de usar esta técnica para la realización de pronósticos, en este caso para temperatura.

La comparación que se observa en la Tabla 1, entre los valores aproximados y los valores reales también se muestran en la Figura 2.

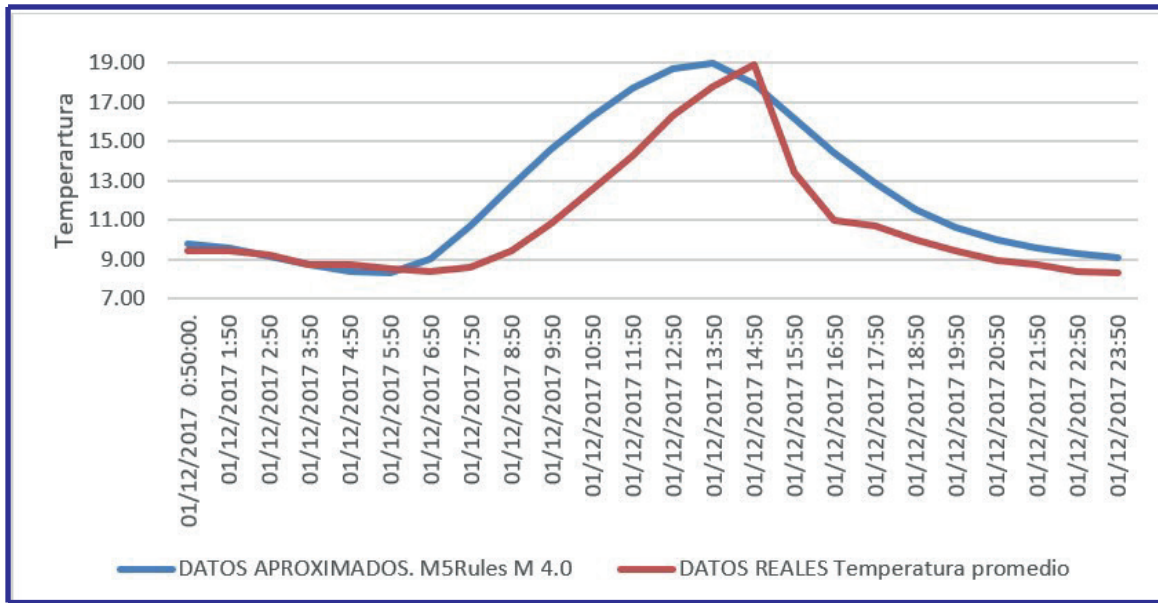
TABLA 1

Comparación de técnicas de pronóstico para temperatura.

Fecha y hora	DATOS REALES		DATOS APROXIMADOS		
	Temperatura promedio	Aprox. M5Rules M 4.0	Aprox. M5P M 4.0	Aprox. Regresión lineal	Aprox. IBk
2017/12/01 00:50:00	9.40	9.7687	9.7687	9.7806	9.9
2017/12/01 01:50:00	9.40	9.5634	9.5634	9.5911	9.3
2017/12/01 02:50:00	9.20	9.157	9.157	9.1951	7.8
2017/12/01 03:50:00	8.70	8.7273	8.7273	8.7731	7.5
2017/12/01 04:50:00	8.70	8.3986	8.3986	8.4501	6.7
2017/12/01 05:50:00	8.50	8.3281	8.3281	8.3858	6.4
2017/12/01 06:50:00	8.40	9.0148	9.0148	9.0768	7
2017/12/01 07:50:00	8.60	10.6923	10.6923	10.7545	11
2017/12/01 08:50:00	9.40	12.7417	12.7417	12.7989	13.4
2017/12/01 09:50:00	10.80	14.6474	14.6474	14.701	16.1
2017/12/01 10:50:00	12.50	16.254	16.254	16.3056	17.8
2017/12/01 11:50:00	14.30	17.7209	17.7209	17.7708	17.8
2017/12/01 12:50:00	16.30	18.6965	18.6965	18.745	20.4
2017/12/01 13:50:00	17.80	18.9718	18.9718	19.0212	20
2017/12/01 14:50:00	18.90	17.9488	17.9488	17.9988	20.3
2017/12/01 15:50:00	13.40	16.169	16.169	16.2185	12.9
2017/12/01 16:50:00	11.00	14.427	14.427	14.4765	9.5
2017/12/01 17:50:00	10.70	12.8479	12.8479	12.8977	10.3
2017/12/01 18:50:00	10.00	11.5147	11.5147	11.5645	10.1
2017/12/01 19:50:00	9.40	10.5999	10.5999	10.6501	9.6
2017/12/01 20:50:00	8.90	10.0119	10.0119	10.0623	9.3
2017/12/01 21:50:00	8.70	9.5711	9.5711	9.6207	9.3
2017/12/01 22:50:00	8.40	9.2683	9.2683	9.3182	9.6
2017/12/01 23:50:00	8.30	9.0782	9.0782	9.1338	9.8
Error cuadrático medio =		2.00	2.00	2.04	2.37

FIGURA 2

Comparativa entre datos reales y aproximados – Temperatura promedio



Con esta verificación se procedió a obtener los pronósticos para los otros parámetros, en cada caso también usando la aplicación Weka. En lo referente a la técnica recomendable a usar ha dependido del parámetro meteorológico a pronosticar.

Humedad relativa promedio

Mínima humedad relativa promedio registrada: 15 %

Máxima humedad relativa promedio registrada: 100 %

Los resultados pronósticos para las 24 horas inmediatas, correspondientes al 01 de diciembre obtenidos con la aplicación Weka y la técnica M5Rules se presentan en la Tabla 2.

TABLA 2

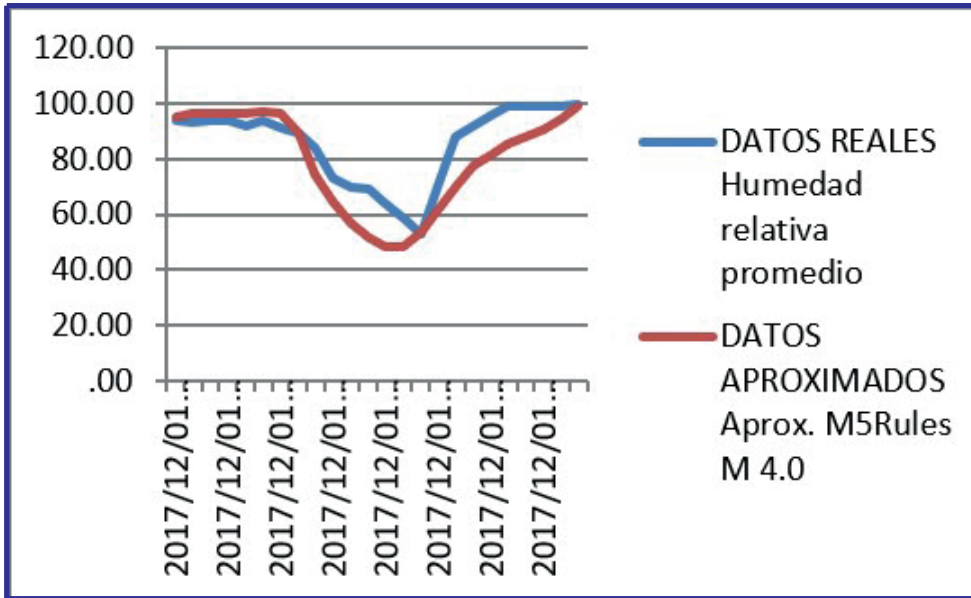
Pronósticos de la Humedad relativa promedio

DATOS REALES		DATOS APROXIMADOS
Fecha y hora	Humedad relativa	
	promedio	Aprox. M5Rules M 4.0
2017/12/01 00:50:00	94.00	95.3921
2017/12/01 01:50:00	93.00	96.1987
2017/12/01 02:50:00	94.00	96.4916
2017/12/01 03:50:00	94.00	96.2446
2017/12/01 04:50:00	92.00	96.6318
2017/12/01 05:50:00	94.00	96.9639
2017/12/01 06:50:00	91.00	96.6287
2017/12/01 07:50:00	89.00	89.3108
2017/12/01 08:50:00	84.00	74.5441
2017/12/01 09:50:00	73.00	64.4385
2017/12/01 10:50:00	70.00	56.8703
2017/12/01 11:50:00	69.00	51.4745
2017/12/01 12:50:00	64.00	48.4024
2017/12/01 13:50:00	59.00	48.1548
2017/12/01 14:50:00	53.00	53.4938
2017/12/01 15:50:00	70.00	61.6183
2017/12/01 16:50:00	88.00	70.0493
2017/12/01 17:50:00	92.00	77.8383
2017/12/01 18:50:00	96.00	81.6234
2017/12/01 19:50:00	99.00	85.2565
2017/12/01 20:50:00	99.00	88.2845
2017/12/01 21:50:00	99.00	90.7637
2017/12/01 22:50:00	99.00	94.5176
2017/12/01 23:50:00	100.00	98.9944

La comparación que se puede observar la Tabla 2, entre los valores aproximados y los valores reales también se muestran en la Figura 3.

FIGURA 3

Comparativa entre datos reales y aproximados – Humedad Relativa



Radiación solar

Mínima radiación solar registrada: 0

Máxima radiación solar registrada: 5063 W/m2

Para el pronóstico de este parámetro meteorológico se verificó que una técnica recomendable es la del vecino más cercano (IBk). Los resultados, que se presentan en

la Tabla 3, se consideran bastante aceptables teniendo en cuenta que la radiación solar puede variar repentinamente por efectos de nubes entre otros motivos. En la Figura 4, muestra la comparativa entre el pronóstico de la radiación solar y los datos reales respectivos.

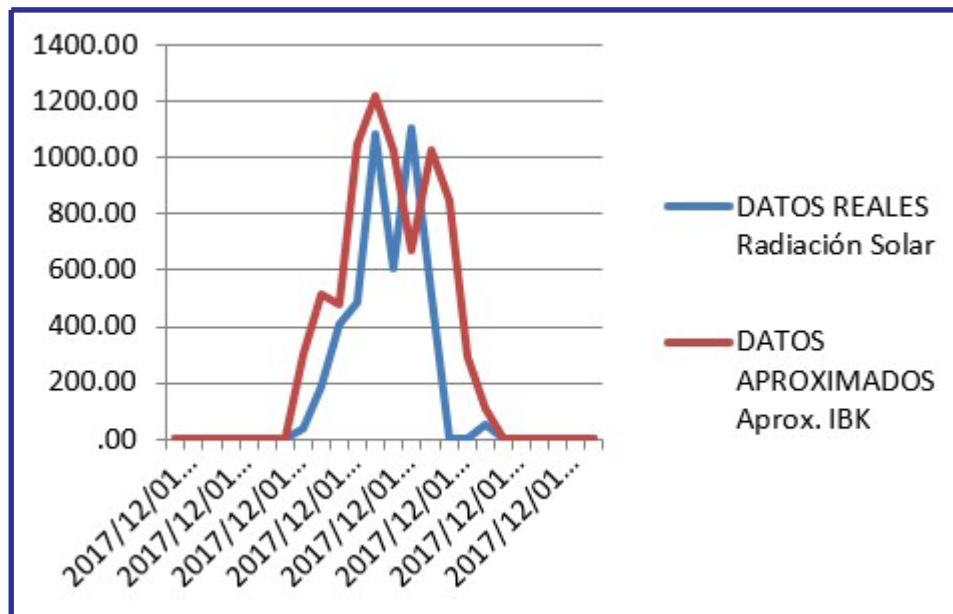
TABLA 3

Pronóstico de radiación solar

DATOS REALES		DATOS APROXIMADOS
Fecha y hora	Radiación Solar	Aprox. IBk
2017/12/01 00:50:00	.00	0
2017/12/01 01:50:00	.00	0
2017/12/01 02:50:00	.00	0
2017/12/01 03:50:00	.00	0
2017/12/01 04:50:00	.00	0
2017/12/01 05:50:00	.00	0
2017/12/01 06:50:00	.00	0
2017/12/01 07:50:00	40.00	298
2017/12/01 08:50:00	189.00	515
2017/12/01 09:50:00	406.00	480
2017/12/01 10:50:00	487.00	1050
2017/12/01 11:50:00	1,082.00	1217
2017/12/01 12:50:00	604.00	1025
2017/12/01 13:50:00	1,103.00	672
2017/12/01 14:50:00	511.00	1025
2017/12/01 15:50:00	.00	851
2017/12/01 16:50:00	.00	297
2017/12/01 17:50:00	52.00	111
2017/12/01 18:50:00	.00	0
2017/12/01 19:50:00	.00	0
2017/12/01 20:50:00	.00	0
2017/12/01 21:50:00	.00	0
2017/12/01 22:50:00	.00	0
2017/12/01 23:50:00	.00	0

FIGURA 4

Comparativa entre datos reales y aproximados – radiación solar

**Punto de rocío promedio por hora**

Mínima temperatura de punto de rocío registrada: - 4.4 °C.

Máxima temperatura de punto de rocío registrada: 12.7 °C.

Para este parámetro meteorológico se verificó también que una técnica excelente para los pronósticos respectivos es la técnica M5Rules.

La Tabla 4, muestra los pronósticos junto a los valores reales, en este caso también para las 24 horas inmediatas (01 de diciembre de 2017) al periodo de referencia (01 de diciembre de 2016 al 30 de noviembre de 2017). El error se ha medido mediante el Error cuadrático medio, que da bastante pequeño: 1.26; por lo que los pronósticos dados son bastante buenos.

TABLA 4
Pronóstico de punto de rocío

DATOS REALES		DATOS APROXIMADOS
Fecha y hora	Punto de rocío	
	promedio	Aprox. M5Rules M 4.0
2017/12/01 00:50:00	9.90	10.3024
2017/12/01 01:50:00	9.70	10.3489
2017/12/01 02:50:00	9.70	10.3472
2017/12/01 03:50:00	9.20	10.4109
2017/12/01 04:50:00	9.00	10.5244
2017/12/01 05:50:00	9.00	10.5178
2017/12/01 06:50:00	8.50	10.5194
2017/12/01 07:50:00	8.40	10.499
2017/12/01 08:50:00	8.30	10.3781
2017/12/01 09:50:00	7.70	10.3779
2017/12/01 10:50:00	8.70	10.4647
2017/12/01 11:50:00	10.10	10.5323
2017/12/01 12:50:00	10.80	10.4787
2017/12/01 13:50:00	11.20	10.3537
2017/12/01 14:50:00	10.50	10.4029
2017/12/01 15:50:00	9.10	10.6931
2017/12/01 16:50:00	10.50	10.9392
2017/12/01 17:50:00	10.90	11.0639
2017/12/01 18:50:00	10.70	11.089
2017/12/01 19:50:00	10.70	11.1213
2017/12/01 20:50:00	10.30	11.0566
2017/12/01 21:50:00	10.00	10.893
2017/12/01 22:50:00	9.80	10.8282
2017/12/01 23:50:00	9.70	10.821
Error cuadrático medio =		1.26

Precipitación total por hora

Mínima precipitación total por hora registrada: 0 mm.

Máxima precipitación total por hora registrada: 147 mm.

El pronóstico de precipitaciones es un problema complejo, como se indica en la tesis de Soto (2013). Debido especialmente a la no periodicidad de las mismas, tanto en fechas como en cantidades de las precipitaciones

tanto en fechas como en cantidades de las precipitaciones. En este trabajo, habiendo visto un problema similar con la radiación solar, se creyó conveniente usar la misma técnica de pronóstico, es decir la técnica del vecino más cercano (IBk). El resultado ha

sido satisfactorio como se puede ver en la Tabla 5 y en la Figura 5, donde se observa la comparación de las poligonales que representan los datos reales y sus aproximaciones.

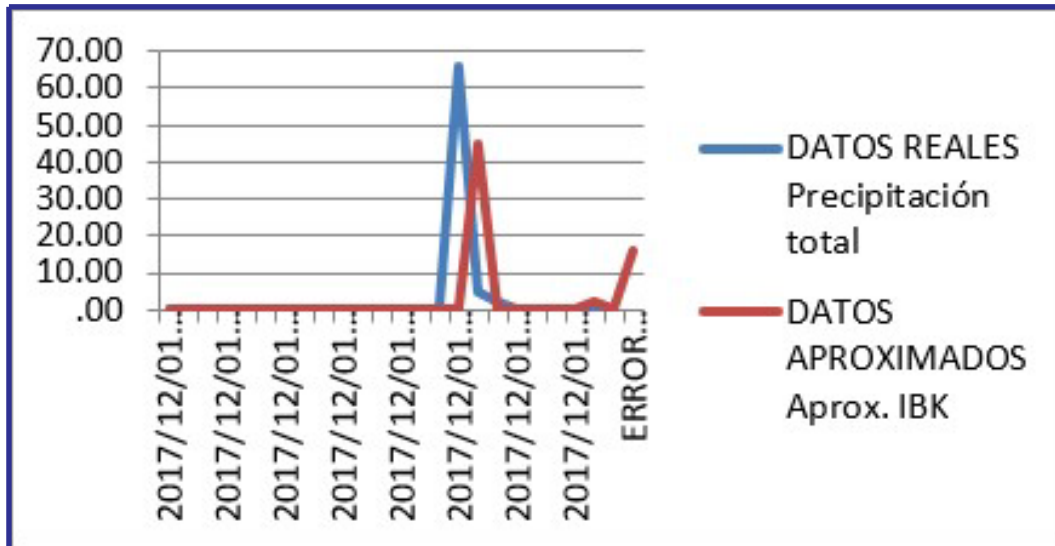
TABLA 5

Pronóstico de precipitación total

Fecha y hora	DATOS	
	DATOS REALES	APROXIMADOS
	Precipitación	
	total	Aprox. IBk
2017/12/01 00:50:00	.00	0
2017/12/01 01:50:00	.00	0
2017/12/01 02:50:00	.00	0
2017/12/01 03:50:00	.00	0
2017/12/01 04:50:00	.00	0
2017/12/01 05:50:00	.00	0
2017/12/01 06:50:00	.00	0
2017/12/01 07:50:00	.00	0
2017/12/01 08:50:00	.00	0
2017/12/01 09:50:00	.00	0
2017/12/01 10:50:00	.00	0
2017/12/01 11:50:00	.00	0
2017/12/01 12:50:00	.00	0
2017/12/01 13:50:00	.00	0
2017/12/01 14:50:00	.00	0
2017/12/01 15:50:00	66.00	0
2017/12/01 16:50:00	5.00	45
2017/12/01 17:50:00	2.00	0
2017/12/01 18:50:00	.00	0
2017/12/01 19:50:00	.00	0
2017/12/01 20:50:00	.00	0
2017/12/01 21:50:00	.00	0
2017/12/01 22:50:00	.00	2
2017/12/01 23:50:00	.00	0

FIGURA 5

Comparativa entre datos reales y aproximados – precipitación total



CONCLUSIONES

- Teniendo en cuenta las evidentes limitaciones de los estudios realizados y su contexto temporal y geográfico, al menos se ha podido comprobar la validez y flexibilidad de las técnicas de minería de datos para tratar de descubrir nuevo conocimiento a partir de datos meteorológicos.
- Se ha mostrado que con la técnica M5Rules se pueden obtener buenos pronósticos para los parámetros temperatura promedio del aire, humedad relativa promedio, así como para punto de rocío.
- También se ha mostrado que, con la técnica del vecino más cercano para pronóstico en minería de datos, se puede obtener buenos pronósticos para radiación solar y precipitaciones.
- Esto dos parámetros tienen la particularidad que van del valor cero a valores más altos en el día y en ciertas estaciones del año, respectivamente.
- Ambas técnicas resultan de sencilla aplicación y permiten obtener modelos de predicción intuitivos e interpretables. Obviamente, como toda técnica de clasificación, la calidad de las predicciones está además condicionada por la calidad del conjunto de datos de entrenamiento.

AGRADECIMIENTOS

Al Centro de Investigación Ambiental para el Desarrollo (CIAD) de la UNASAM, en Huaraz - Perú por facilitar la información meteorológica necesaria para este trabajo.

REFERENCIAS BIBLIOGRAFICAS

- Abraham, A., Grosan, C., & Ramos, V. (2006). *Swarm Intelligence in Data Mining*. Springer. <https://link.springer.com/book/10.1007/978-3-540-34956-3>
- Duggal, H., & Singh, P. (2012). Comparative Study of the Performance of M5-Rules. *Journal of Software Engineering and Applications*, 5(4), 270 - 276. DOI: 10.4236/jsea.2012.54032
- Ecured. (s.f.). *Ecured*. <https://www.ecured.cu/Weka>
- Eraso, Y. (2020). *Comportamiento de las variables meteorológicas en el periodo 2021 al 2020 en el Valle del Mantaro*. [Tesis de Licenciatura. Universidad Continental]. https://repositorio.continental.edu.pe/bitstream/20.500.12394/9078/4/IV_FIN_107_TI_Eraso_Tiza_2020.pdf
- Espinola, J., Pla, L., Montanez, E., Leyva, J., & Cáceres, V. (2017). Evaluación de sustentabilidad del sistema agrícola de la Comunidad de Huapra (Perú). *Revista Investigación Operacional*, 38(1), 91 - 100. <https://go.gale.com/ps/i.do?p=IFME&u=googleScholar&id=GALE|A487601231&v=2.1&it=r&sid=googleScholar&asid=cfa44a78>
- Haro, S., Zíñiga, L., Meneses, A., Vera, L., & Escudero, A. (2018). Métodos de clasificación en minería de datos meteorológicos. *Perfiles - Revista científica*, 20(2), 107 - 113. http://dspace.espoeh.edu.ec/bitstream/123456789/9395/1/per_n20_v2_13.pdf
- Hernández, J. (2004). *Introducción a la Minería de Datos*. Pearson Educacion S.A. [https://www.pearsoneducacion.net/peru/Inicio/introduccion-a-la-mineria-de-datos-\(ebook\)](https://www.pearsoneducacion.net/peru/Inicio/introduccion-a-la-mineria-de-datos-(ebook))
- Hitachi Vantara. (2022). *Pentaho Community Edition*. <https://www.hitachivantara.com/en-us/pdf/white-paper/pentaho-community-edition-installation-guide-for-windows-whitepaper.pdf>
- Martínez-Abad, F., & Hernández-Ramos, J. (2018). Técnicas de minería de datos con software libre para la detección de factores asociados al rendimiento. *REXE - Revista de Estudios y Experiencias en Educación*, 2(2), 135 - 145. <https://www.redalyc.org/journal/2431/243156768012/html/>
- Rey, T., Kordon, A., & Wells, C. (2012). *Applied Data Mining for Forecasting Using SAS*. SAS Institute. https://support.sas.com/content/dam/SAS/support/en/books/applied-data-mining-for-forecasting/63533_excerpt.pdf
- Servicio Nacional de Meteorología e Hidrología del Perú [SENAMHI] (s.f.). *Servicio Nacional de Meteorología e Hidrología del Perú*. <https://www.gob.pe/senamhi>
- Soto, C. (2013). *Minería de datos para series temporales en Weka y su aplicación en el pronóstico de precipitaciones*. [Tesis de Licenciatura. Universidad Central "Marta Abreu" de Las Villas]. <https://dspace.uclv.edu.cu/bitstream/handle/123456789/1935/Miner%C3%ADa%20de%20datos>

%20para%20series%20temporales%20en%20
Weka%20y%20su%20aplicaci%C3%B3n%20en
%20el%20pron%C3%B3stico%20de%20precipit
aciones.pdf?sequence=1&isAllowed=y

Stratebi (s.f.). *Stratebi - open business intelligence*.
<http://www.stratebi.com/pentaho>

Weka. (s.f.). *Weka*.
[https://waikato.github.io/weka-wiki/downloading_
weka/](https://waikato.github.io/weka-wiki/downloading_weka/)

Witten, H., Frank, E., Hall, M., & Pal, C. (2016). *Data Mining: Practical Machine Learning Tools and Techniques (4th ed.)*. Morgan Kaufmann.
[https://libgen.is/book/index.php?md5=2FF05CC2
C0F0B0C6D857052092F0E16E](https://libgen.is/book/index.php?md5=2FF05CC2C0F0B0C6D857052092F0E16E)