

Trabajo Número 02

Estudiante: Jhon Gesell Villanueva Portella

Materia: Tópicos Avanzados en Ingeniería de Software

A. Objetivo: Aplicar regresión lineal mediante el uso de workflow mediante Knime desde la plataforma Windows 10.

Procedimiento:

01. Descarga de la data desde la URL:

<https://www.datosabiertos.gob.pe/dataset/temperatura-agua-y-ambiente-nodo-sensor-puerto-maldonado-represa-iiap>

Trabajamos con variables de de temperatura de la superficie del agua así como del aire, el archivo de extensión .xlsx es descargado y lo transformamos a un archivo de extensión .csv por comodidad en la lectura con Knime , habiendo ubicado ambas columnas de datos en una hoja del archivo como se muestra en la captura de pantalla de la figura 01.

	A	B	C
1	Mon Fecha	TempAgua	TempAire
2	10/11/2016	28.637142	24.957142
3	11/11/2016	29.275854	27.705454
4	14/11/2016	29.72729	27.735483
5	15/11/2016	28.954698	27.047413
6	16/11/2016	29.753428	27.48
7	17/11/2016	29.611875	28.142045
8	18/11/2016	31.379444	28.222222
9	19/11/2016	30.239494	26.60606
10	20/11/2016	29.315	30.638297
11	21/11/2016	28.637102	27.267045
12	22/11/2016	28.323936	24.468085
13	23/11/2016	28.079703	25.725925
14	24/11/2016	27.435703	25.992592
15	25/11/2016	27.505257	26.097938
16	26/11/2016	29.218385	28.666666
17	27/11/2016	28.5918	26.466666
18	28/11/2016	27.944081	24.265306
19	29/11/2016	31.262719	31.122807
20	30/11/2016	27.54725	14.975
21	01/12/2016	27.666	32.6
22	02/12/2016	27.389411	25.372549
23	03/12/2016	29.35884	25.260869
24	04/12/2016	28.532829	16.24031
25	07/12/2016	29.182792	28.51948
26	08/12/2016	28.614867	26.693121
27	09/12/2016	29.754477	28.20398

Fig. 01

02. Cargamos el Knime y creamos un nuevo entorno de trabajo, señalamos el flujo de trabajo:

02.01. Para la lectura , ubicamos el archivo de extensión CSV y ejecutamos, quedando procesado.

02.02. Arrastramos del 'Node Repository' del software y nos quedamos con las variables de interés, quedando como en la Fig. 02, seguido picamos en el botón sobre el nodo la opción 'Execute and OpenView', mostrando la fig. 03.

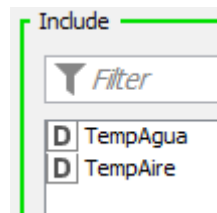


Fig. 02

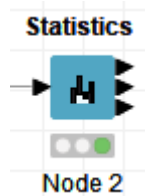


Fig. 03

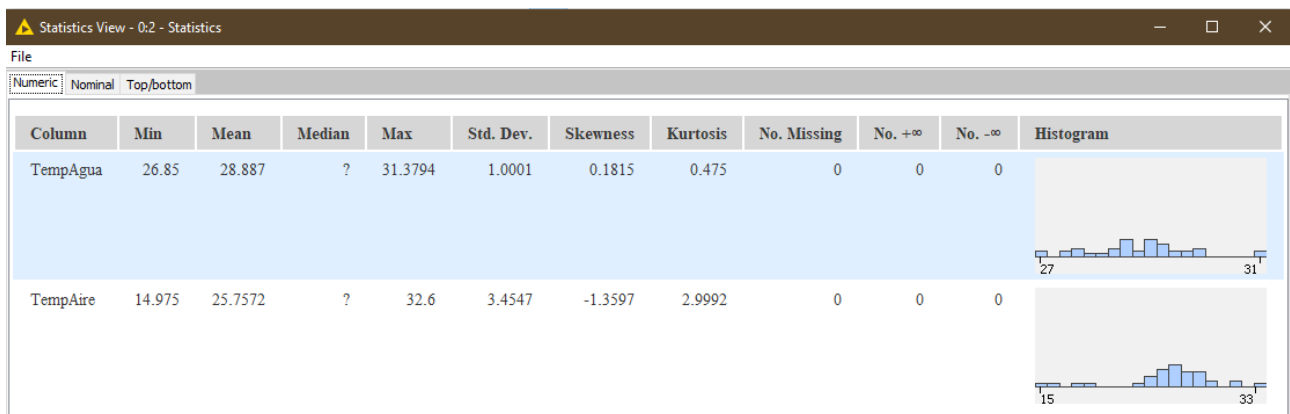


Fig. 04

02.03. Arrastramos y ejecutamos el nodo Scatter Plot (Local), ejecutamos y visualizamos la salida para las variables con las que estamos trabajando que son temperatura para la superficie del agua y del aire, quedandonos como resultado la fig. 05.

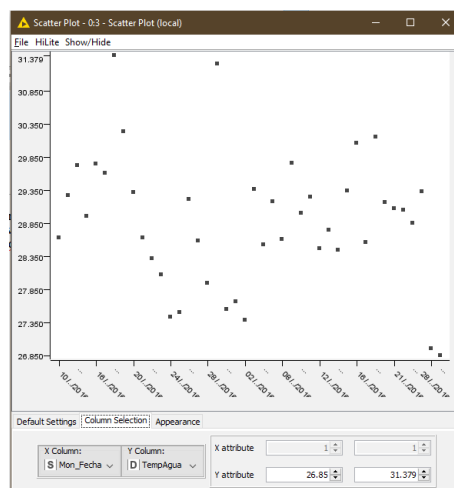


Fig. 05

02.04. Arrastramos el nodo Linear Regression Learner, ejecutamos y visualizamos quedando como resultado la fig. 06. para la selección de la temperatura del agua. La fig. 07 nos da la tabla de resultados para este caso.

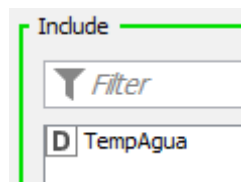


Fig. 06

Linear Regression Resu...

File

Statistics on Linear Regression

Variable	Coeff.	Std. Err.	t-value	P> t
TempAgua	1.5347	0.4893	3.1364	0.0032
Intercept	-18.5765	14.1436	-1.3134	0.1965

Multiple R-Squared: 0.1974
Adjusted R-Squared: 0.1773

Fig. 07

02.05. Esquema del diagrama final conectado en Knime.

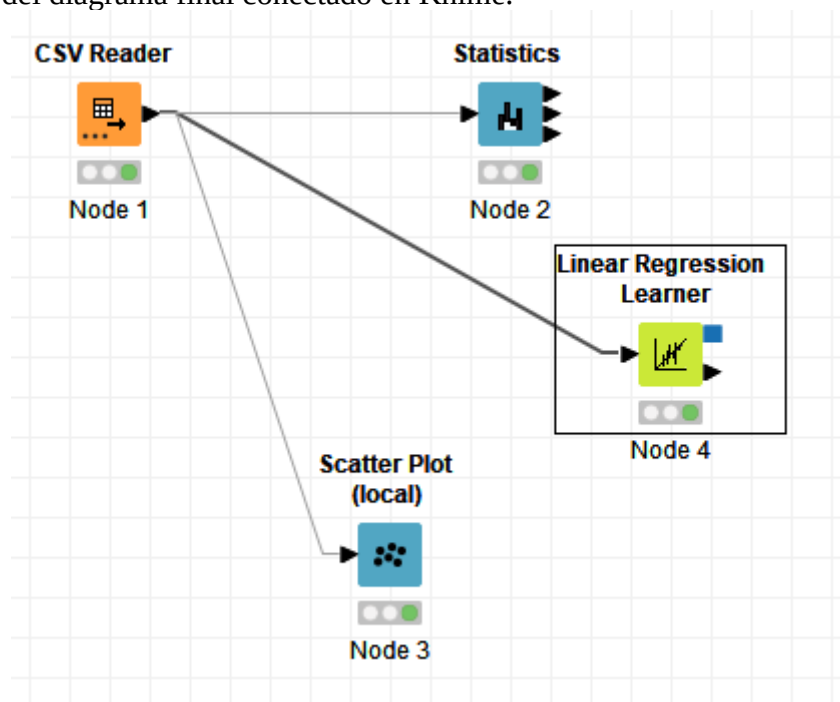


Fig. 08

Conclusiones: Luego de haber iniciado el procedimiento para lograr obtener resultados de una regresión lineal mediante el software Knime , al ver la relación que existe entre los datos de la temperatura superficial del agua y del aire , no se distinguía exactamente una tendencia linea de los datos.

El modelo de la regresión lineal se ejecutó aunque no es el más adecuado para describir la tendencia de la mejor forma correcta para este datasets según el valor del R2 de la fig. 07.