



DetECCIÓN de anomalías en la industria

Héctor Fornes Gabaldón (autor)

Carlos David Martínez Hinarejos (tutor)

Lorenzo Martínez Moret (cotutor en Nunsys)





Índice

01 Introducción

Introducción y objetivos

02 Conceptos

Definición de anomalía, LSTM, Autoencoder, GAN

03 *Dataset*

Descripción de los datos

04 Implementación

Arquitecturas usadas y detalles de implementación

05 Experimentación

Experimentación y evaluación

06 Conclusiones

Conclusiones y trabajos futuros



01

Introducción y objetivos



Objetivos

Cliente

Satisfacer sus necesidades de detección de anomalías

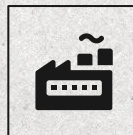


Implementación

De las técnicas más prometedoras

Estudio

De diferentes algoritmos que representan el estado del arte



Despliegue

Implantar la solución en las instalaciones del cliente





02

Conceptos fundamentales





Definición de anomalía

Definición

“Un dato anómalo es la consecuencia medible de un cambio de estado inesperado de un sistema, que se sale fuera de su norma local o global.”

A. A. Cook et al.

Tipos

- Puntual
- Contextual
- Colectiva

Desafíos

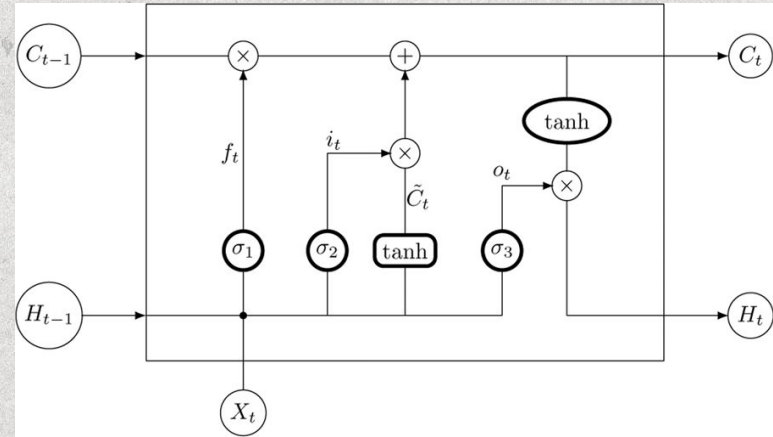
- Naturaleza de las series de datos temporales
- Escasez de datos anómalos
- Restricciones de tiempo y recursos

Redes LSTM

RNN que retiene dependencias de largo alcance en series temporales

Puertas:

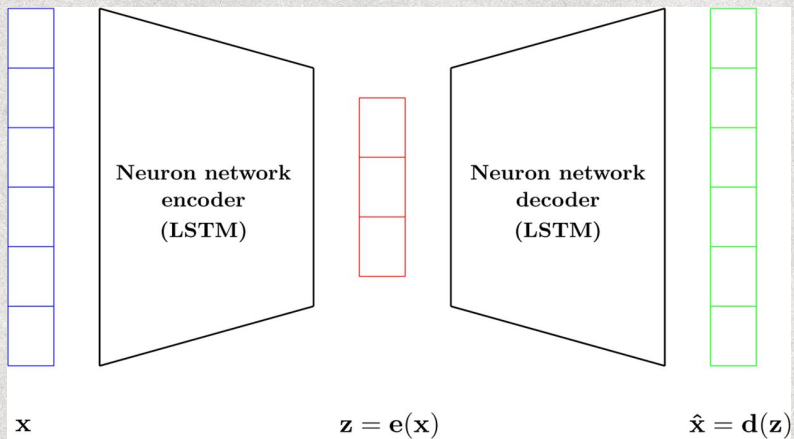
- De olvido
- De entrada
- De salida



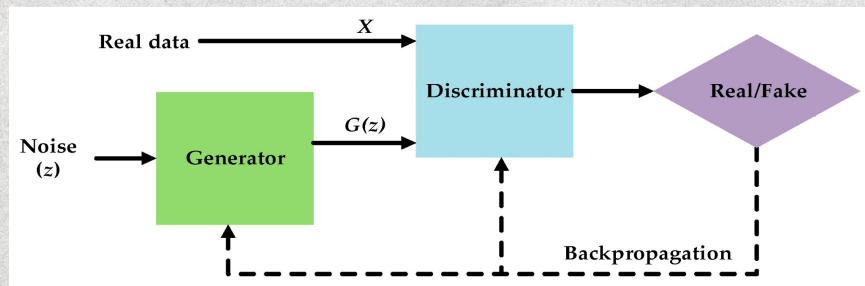


Autoencoder y GAN

Autoencoder



Generative Adversarial Networks





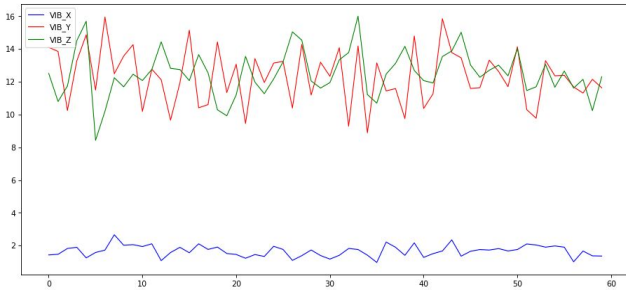
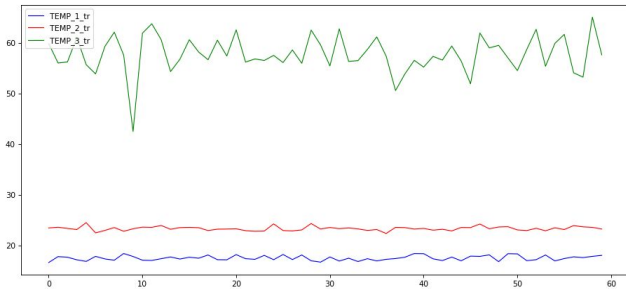
03

Descripción de los datos

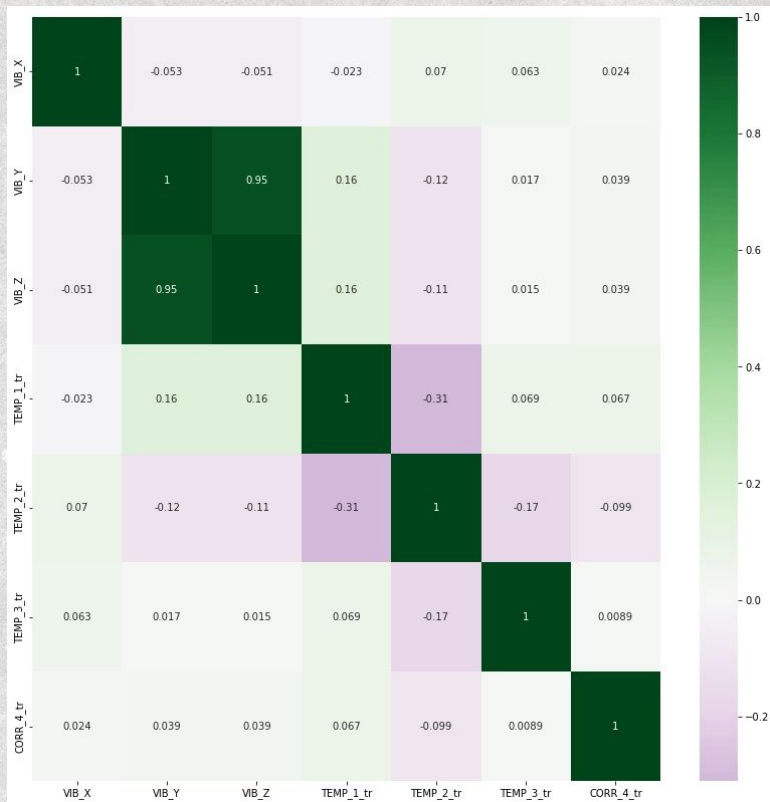




Descripción de variables

Variables	Tipo	Gráficos
<ul style="list-style-type: none">VIB_XVIB_YVIB_Z	Vibración	
	Medida en los 3 ejes	
<ul style="list-style-type: none">TEMP_1_trTEMP_2_trTEMP_3_tr	Temperatura	
	Escala de 0 a 100	
<ul style="list-style-type: none">CORR_4_tr	Corriente	
	Escala de 0 a 10	

Coeficiente de correlación de *Pearson*



VIB

Ejes Y y Z
relacionados,
pero no X

TEMP

Canales 1 y 2
con correlación
negativa

CORR

Sin correlación
con otras
variables



04

Implementación





Arquitecturas usadas

LSTM *Autoencoder*

- *Encoder*
- 3 capas BiLSTM
- 64, 32 y 16 neuronas
- *Decoder*: opuesta

GAN

- **Generador**: como LSTM *Autoencoder*
- **Función crítica Cx**: real vs sintético
- **Función crítica Cz**: eficiencia de codificación
- **Colapso de modo**: pérdida *Wassertein*



Detalles de implementación

Obtención de puntuaciones de anomalía

- Errores de reconstrucción: diferencia punto a punto, de área y DTW
- Salida de Cx
- Combinación

Umbrales de error dinámicos

- Método no supervisado
- Suavizado del error
- Ventanas deslizantes: ¿cuántos *scores* antiguos influyen al calcular el umbral actual?

Mitigación de falsos positivos

- Exceso de falsos positivos
- Procedimiento de poda
- Umbral mínimo



05

Experimentación y evaluación

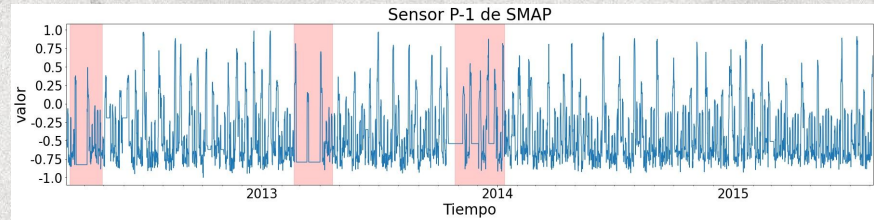


Dataset SMAP

Los datos del cliente carecen de anomalías.

Pruebas con *dataset* SMAP:

- Datos satélite NASA
- Se han usado 7 sensores
- Anomalías en 2012, 2013 y 2014



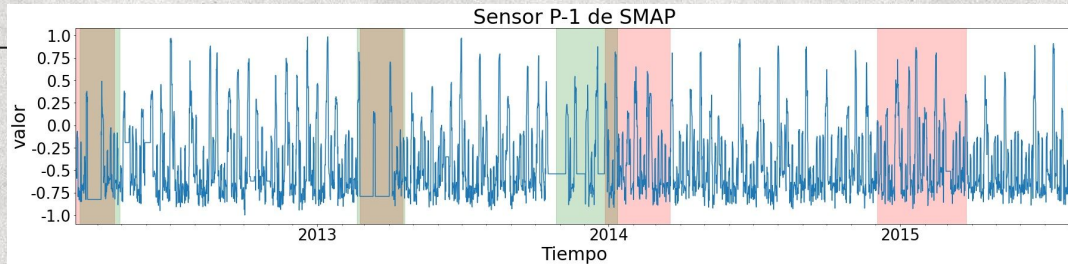


Parámetros

- ***Epochs:***
 - 50, 100 y 200
- ***Learning rate:***
 - LSTM *Autoencoder*: 0.0001 y 0.0005
 - GAN: 0.0005 y 0.001
- **Métodos de cálculo de error:**
 - Diferencia punto a punto, de área y DTW
- **Combinación de puntuaciones GAN:**
 - Suma y multiplicación
- **Tamaños de ventana y de avance de ventana en cada paso:**
 - *Window size*: 0.2, 0.33, 0.5 y 0.8
 - *Window step size*: 0.1, 0.2, 0.33 y 0.5



Método de evaluación



Valor F1

Media armónica entre la precisión y la sensibilidad

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$



Resultados GAN

epochs/lr	0.0001	0.0005
50	0.50	0.67
100	0.57	0.86
200	-	0.25

Resultados obtenidos con diferencia de área, combinación por multiplicación, WS de 0.33 y WSS de 0.33



Resultados LSTM *Autoencoder*

	<i>learning rate - epochs</i>			
	0.0005 - 50	0.001 - 50	0.0005 - 100	0.001 - 100
WSS / WS				
0.20 / 0.50	0.57	0.22	0.22	0.33
0.50 / 0.20	0.50	0.25	0.20	0.44

Resultados obtenidos con diferencia de área



06

Conclusiones y trabajos futuros



Conclusiones y trabajos futuros

Estado actual

Fase de pruebas en instalaciones del cliente



Transferencia

Transfer learning: DDC, DAN, Deep CORAL

Atención

Mecanismos de atención



Predicción

Mantenimiento predictivo





Gracias por su atención

hector.fornes@nunsys.com

Nunsys SA

