

# **TÉCNICA DEL APRENDIZAJE NO SUPERVISADO: DETECCIÓN DE ANOMALÍAS**

Nelson Yoel Phuño Cahuana; Daysi Saimira Machaca Condori

## **INTELIGENCIA COMPUTACIONAL**

[nelsonyoelpc@gmail.com](mailto:nelsonyoelpc@gmail.com); [zaimira.mlc@gmail.com](mailto:zaimira.mlc@gmail.com)

### **DETECCION DE ANOMALÍAS**

El objetivo es conocer cómo se ven los datos normales y se usa para detectar instancias anormales, como artículos defectuosos en una producción, línea o una nueva tendencia en una serie de tiempo. La detección de anomalías es una técnica estadística que usa Analytics Intelligence para identificar valores atípicos tanto en datos de series temporales de una métrica determinada como dentro de un segmento en un mismo momento determinado (Géron, 2019).

¿Cómo funciona?

La detección de anomalías (outliers) con PCA y Autoencoders es una estrategia no supervisada para identificar anomalías cuando los datos no están etiquetados, es decir, no se conoce la clasificación real (anomalía - no anomalía) de las observaciones. No utiliza directamente su resultado como forma de detectar anomalías, sino que emplea el error de reconstrucción producido al revertir la reducción de dimensionalidad. El error de reconstrucción como estrategia para detectar anomalías se basa en la siguiente idea: los métodos de reducción de dimensionalidad permiten proyectar las observaciones en un espacio de menor dimensión que el espacio original, a la vez que tratan de conservar la mayor información posible. La forma en que consiguen minimizar la pérdida global de información es buscando un nuevo espacio en el que la mayoría de observaciones puedan ser bien representadas. (Rodrigo, 2020)

¿Cómo se mide su eficiencia?

Según la detección de anomalías basada en series temporales utiliza el historial de datos para marcar una única métrica dentro de un valor de dimensión, también ofrece un sistema de detección simultánea de anomalías en varios valores de dimensiones y métricas en un momento dado. En este sistema utiliza el análisis de componentes

principales (ACP) para aprovechar la estructura de correlación de las métricas junto con la validación cruzada para marcar anomalías.

¿A qué tipos de problema mejor responde?

- Señales de electro cardiogramas de la imagen.
- Detección de fraude en Tarjetas de Crédito
- Monitoreo de la Seguridad en la Red
- Monitoreo del ritmo cardiaco
- Detectar brotes de epidemias

## **Bibliografía**

ANDRÉS J. HERNÁNDEZ, E. D. (2006). REDUCCIÓN DE DIMENSIONES PARA CLASIFICACIÓN DE DATOS MULTIDIMENSIONALES USANDO MEDIDAS DE INFORMACIÓN . (32).

Géron, A. (2019). Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow. Nicole Tache.

Rodrigo, J. A. (01 de 02 de 2020). Recuperado el 15 de 02 de 2021, de [https://www.cienciadedatos.net/documentos/52\\_deteccion\\_anomalias\\_autoencoder\\_pca.html#:~:text=La%20detecci%C3%B3n%20de%20anomal%C3%ADas%20\(outliers,no%20anomal%C3%ADa\)%20de%20las%20observaciones.](https://www.cienciadedatos.net/documentos/52_deteccion_anomalias_autoencoder_pca.html#:~:text=La%20detecci%C3%B3n%20de%20anomal%C3%ADas%20(outliers,no%20anomal%C3%ADa)%20de%20las%20observaciones.)