



Módulo II. Métodos de Machine Learning

CURSO:

**Taller de Machine Learning
para el análisis y visualización
en Power BI**

Laboratorio de Inteligencia Artificial

Tema 05 – Casos de estudio

Profesor: Saúl Domínguez Isidro, PhD.
Contacto: saul.dominguez@lania.edu.mx

Contenido

- Modelo de predicción de generación energía
- Modelo de predicción de fallas

Caso 1

- Una central eléctrica de ciclo combinado (CCPP) está compuesta por:
 - turbinas de gas (GT),
 - turbinas de vapor (ST)
 - y generadores de vapor de recuperación de calor.
- En un CCPP, la electricidad se genera mediante turbinas de gas y vapor, que se combinan en un ciclo y se transfieren de una turbina a otra. Si bien el vacío generado por ST tiene una influencia, otras tres variables ambientales afectan el rendimiento del GT.

Propiedades

- El conjunto de datos contiene 9568 puntos de datos recopilados de una central eléctrica de ciclo combinado durante 6 años (2006-2011), cuando la central eléctrica estaba configurada para funcionar con carga completa.
- Las características consisten en las variables ambientales promedio por hora:
 - Temperatura (AT),
 - Presión ambiente (AP),
 - Humedad relativa (RH)
 - Vacío de escape (V)
 - Predecir la producción neta de energía eléctrica (EP) por hora de la planta. (Target)

Caso 2

- Una empresa manufacturera, genera sus productos mediante diversas máquinas. Como la producción se detiene cada vez que una máquina tiene un error, la administración desea crear un modelo predictivo que encuentre qué máquina va a fallar a continuación.
- La empresa está utilizando 1000 máquinas.
 - En promedio, estas máquinas tienen un fallo cada 55 semanas.
 - Algunas de estas máquinas son completamente nuevas, otras han estado funcionando durante casi dos años.
 - En nuestro conjunto de datos, casi el 40 % de las máquinas tuvieron un fallo en los últimos dos años

Propiedades

- Tipo de tarea: clasificación binaria
- Número de columnas: 9
- Número de filas: 1000 muestras
- Target: (broken) ¿Máquina realmente rota? Si/No.
- Balance: clase positiva (broken) 40%, clase negativa: 60%

Variables

- Aleatorio (random)
- Número de máquina (Machine nbr): de 1 a 1000
- Vida útil (lifetime): indica el número de semanas desde que se ha utilizado la máquina
- Variables numéricas:
 - Temperatura (Temperature)
 - Presión (Pressure)
 - Humedad (Moisture)
- El equipo que usa la máquina (team)
- El proveedor de la máquina (provider)
- "broken", que es nuestro taget (sí o no)