

## Actividad. Inteligencia Artificial

Unidad	6
Entrega	Proyecto de Machine Learning

### 1. Enunciado

Nuestro cliente nos ha solicitado realizar un proyecto piloto para el uso de técnicas de inteligencia artificial en su fábrica. Nos solicita llevar a cabo un estudio para identificar qué variables van a ser necesarias para el entrenamiento de los modelos de aprendizaje automático (ML).

Quiere que resolvamos dos problemas:

1. Primero, quiere identificar que a qué línea pertenece los datos que recoge, ya que llegan a un sistema central pero no sabe de donde provienen los datos. He realizado una asignación manual de parte de los datos recogidos, pero el proceso de asignación manual es muy costoso como para mantenerlo en el tiempo. Por lo que quiere que entrenemos un modelo para clasificar los datos en función de la línea de la que provienen.
2. Segundo, quiere eliminar los medidores del gramaje del papel filtrante que produce ya que requieren un alto mantenimiento. Por lo que deberemos entrenar un modelo de regresión para estimar cual va a ser el gramaje del papel producido.

Quiere que se evalúe el rendimiento de los modelos con al menos 2 métricas distintas.

Recordamos las características de las plantas de producción de nuestro cliente:

En esta planta se fabrican y ensamblan varios componentes para su posterior distribución. La planta cuenta con varias líneas de producción en las que intervienen una o más estaciones de trabajo.

En función de las tareas realizadas a lo largo del proceso productivo, de la maquinaria que intervenga y de los componentes producidos se realizan mediciones por distintos sensores. Hay dos líneas de producción donde se producen unos filtros que posteriormente son ensamblados en otros componentes.

Las líneas 400 y 500 son las encargadas de realizar un proceso de corte de un papel filtrante por ultrasonidos para su posterior incorporación a las líneas 4 y 5 donde se suelda para terminar de ensamblar los filtros. Los datos que contiene el fichero son los siguientes:

1. LOCATION\_ID: Identificador único para la localización dentro de las distintas plantas de producción.
2. Line: Línea de producción (para un mismo producto puede disponerse de varias líneas para aumentar el volumen de producción).
3. Station: Estación física de trabajo diseñada para realizar uno o varios procesos productivos (una estación puede contar con varios puestos de trabajo). La etiqueta indica un tipo de estación no una unidad física.
4. Index: Indica el número de estaciones del mismo tipo que hay dentro de la línea. Ej. Línea: 400, Station:20, Index:1,2 (si hay dos estaciones de tipo 20 dentro de la línea 400).
5. Fu: Function Unit. Separaciones funcionales dentro de la estación e index. Cuando dentro de una estación se realizan dos funciones distintas.
6. WP: Work Position. Posición donde se pueden realizar uno o más procesos, dentro de una function unit (ej. envasar y estampar).
7. TP: Tool Position. Posición de la herramienta dentro del WP. Varía cuando tienes varias herramientas involucradas en la misma WP.
8. RESULT\_DATE: Fecha, hora y franja horaria.

## Actividad. Inteligencia Artificial

9. UNIQUEPART\_ID: Identificador único de cada una de las piezas producidas.
10. Cut Amplitude: Amplitud del corte realizado (micrómetros).
11. Cut Energy: Energía utilizada en el corte (J).
12. Cut Frequency: Frecuencia ultrasonido utilizada para realizar el corte (Hz).
13. Cut Power: Potencia de corte (W).
14. Cut WeldTime: Tiempo necesario para realizar el corte del papel (ms).
15. Grammage: Gramaje del papel de lija (g/m<sup>2</sup>).
16. Humidity: Humedad del entorno (%).
17. Temperature: Grados del entorno.

Nuestro cliente nos pide lo siguiente:

1. Utilizar uno o dos métodos de selección de variables. Evaluar el resultado y exponer los métodos utilizados. Posteriormente comunicar las variables que hemos decidido descartar (para cada uno de los dos problemas que vamos a resolver).
2. Entrenar un modelo de ML para realizar tareas de clasificación. Dicha tarea consiste en determinar qué línea de producción es la que está generando los datos.
3. Entrenar un modelo de ML para realizar tareas de regresión. Dicha tarea consiste en determinar un valor de gramaje del papel filtrante a partir del resto de datos recogidos.
4. Utilizar las métricas expuestas en el ejercicio para evaluar el funcionamiento de los modelos generados. Añadir y explicar dos métricas nuevas para la evaluación de los modelos generados (2 para el clasificador y 2 para el regresor).

Además, nuestro cliente nos ha hablado sobre lo útil que podría ser estimar el valor de la frecuencia de corte:

5. Entrenar un modelo de ML complejo (ej. Random Forest, NN...) y realizar un proceso de ajuste de parámetros (sin este proceso no será evaluada esta parte). El problema que se quiere resolver es la estimación del valor de la frecuencia de corte utilizada. Exponer el rendimiento final del modelo utilizando las métricas que se consideren oportunas, describir los parámetros del algoritmo y justificar el ajuste de parámetros realizado.

## 2. Detalles de la entrega

1. Elabora un documento que incluya la siguiente información:
  - a. Estudios y análisis realizados para el cliente, junto con los resultados obtenidos por los modelos entrenados. Esta información debe estar soportada por gráficos.
  - b. Código por el cual han sido entrenados los modelos (archivos. ipynb o .py).
2. Sube de forma individual el documento y cualquier información anexa que se haya mencionado en el documento y se considere relevante.



### Importante

Se valorará cualquier información que se crea interesante.

## Actividad. Inteligencia Artificial

Cualquier duda no dudéis en preguntarme. Si el ejercicio resulta muy difícil podemos hacer una clase virtual donde veamos algunos de los puntos.

---

### 3. Anexo

Tener en cuenta que se necesitan los siguientes archivos:

1. CleanedData.csv
2. ML\_A6.ipynb