**Instituto Tecnológico y de Estudios Superiores de Monterrey**

Logotipo, nombre de la empresa

Descripción generada automáticamente

Machine Learning para Secadoras

**Materia:**

Inteligencia Artificial Avanzada

Integrantes:

José Antonio Miranda Baños

**Campus:** Querétaro.

**Fecha: 24 de agosto de 2024**

Machine Learning Para Secadoras

José Antonio Miranda Baños

*IRS, Instituto Tecnológico y de Estudios Superiores de Monterrey   
(Campus QRT)*Qrt, Mx  
*A0161795@tec.mx*

*Abstract*—

Keywords— Machine Learnig, Regresión logistica,

# Introducción

Esta década (2020) ha sido definitivamente la década del Machine Learning; pareciera que todas las empresas buscan implementar esta tecnología en cada uno de sus productos. El Machine Learning ha demostrado ser una herramienta poderosa en diversas industrias, desde la automotriz hasta la salud. En el contexto de los electrodomésticos, particularmente en las secadoras, el uso de Machine Learning abre nuevas posibilidades para mejorar la eficiencia, optimizar el rendimiento y personalizar la experiencia del usuario.

Las secadoras tradicionales han evolucionado en términos de diseño y funcionalidad, pero la integración de algoritmos de Machine Learning promete llevar estas mejoras a un nuevo nivel. Con la capacidad de analizar grandes volúmenes de datos en tiempo real, las secadoras equipadas con Machine Learning podrán predecir patrones de uso, ajustar los ciclos de secado según las necesidades específicas de cada carga y anticipar problemas potenciales antes de que ocurran.

De modo que es solo cuestión de tiempo para empezar a ver más productos del tipo electrodoméstico con integraciones de Machine Learning

Este artículo explora cómo el Machine Learning puede transformar las secadoras, revisando aplicaciones actuales, desafíos técnicos y perspectivas futuras. Además, se discute cómo la implementación de estas tecnologías no solo puede incrementar la eficiencia energética, sino también ofrecer una mayor comodidad y cuidado para las prendas. A través de un análisis detallado de estudios de caso y avances recientes, se busca proporcionar una visión integral de cómo el Machine Learning está redefiniendo el sector de los electrodomésticos.

# desarrollo

## ¿Como funciona tu Secadora?

Aunque la pregunta puede parecer sencilla, entender el funcionamiento de una secadora es crucial para abordar cuestiones más profundas como el daño potencial a la ropa o la eficiencia en el manejo de grandes volúmenes de prendas. Las secadoras, a pesar de su diseño aparentemente simple, involucran una serie de procesos complejos que afectan tanto el resultado del secado como la durabilidad de las prendas.

El funcionamiento de una secadora es relativamente sencillo y se puede entender cómo que queremos separar el agua que quedo después de que la lavadora centrifugara nuestra ropa. La mejor solución para esto es calentarla para evaporar esta agua y extraer el vapor para ir lentamente reduciendo la cantidad de agua en nuestra ropa. Ahora hay muchas otras cosas involucradas que no considere en el resumen rápido de como secar ropa. Por ejemplo, no tomamos en cuenta que la ropa no puede estar junta o sino esta se acumula y esta no se llega a secar, que no podemos hacer girar el tambor muy rápido sino el aire no alcanza a llevarse nada de la humedad y la obstrucción del escape de tu secadora.

Estos y otros varios factores llegan a afectar el ciclo de secado de tu secadora y dañar tu ropa, por lo que no saber usar bien tu secadora afectara la vida útil de tus prendas.

## Machine Learning para Secadoras

*Al comenzar este proyecto, no anticipé la magnitud del impacto que el Machine Learning (ML) podría tener en las secadoras. Inicialmente, podría parecer que no hay necesidad de implementar modelos de ML en un electrodoméstico como una secadora, especialmente si consideramos que, a diferencia de una estufa, no estamos presentes para supervisar su funcionamiento. Sin embargo, al profundizar en el análisis, se vuelve evidente que las secadoras, al ser herramientas "Non-Supervised", se benefician enormemente de la integración de modelos de ML.*

1. *Definición de "Non-Supervised" y "Supervised"*

*Non-Supervised: En el contexto de los electrodomésticos, una máquina "Non-Supervised" es aquella que realiza una tarea de manera autónoma sin la necesidad de supervisión continua. Un ejemplo de esto es la secadora. Una vez que introduces la ropa y seleccionas un programa, la secadora completa el ciclo sin intervención adicional. Esta falta de supervisión directa la hace candidata ideal para la integración de ML, que puede optimizar su funcionamiento y mejorar los resultados sin la necesidad de monitoreo constante.*

*Supervised: En contraste, herramientas "Supervised" como la estufa requieren la presencia del usuario para ajustar parámetros en tiempo real. Dado que el usuario está activamente involucrado en el proceso (como cocinar), no hay una necesidad inmediata de implementar ML para gestionar la operación del aparato, ya que el control manual y la supervisión directa cumplen esa función.*

# Implementación

El desafío es claro: desarrollar la secadora ideal, una máquina que pueda secar eficientemente cualquier tipo de ropa sin importar su naturaleza, al punto de que entregar la ropa doblada sea casi una muestra de vanidad. Para alcanzar este objetivo, nos proponemos utilizar técnicas avanzadas de Machine Learning, específicamente una regresión logística, para identificar y ajustar el ciclo de secado adecuado para cada tipo de prenda.

##### Definiendo la Secadora Ideal

El concepto de la secadora ideal implica una máquina capaz de adaptar su ciclo de secado a las características específicas de cada carga de ropa, asegurando resultados óptimos sin necesidad de intervención manual. Esto requiere un sistema que pueda:

Detectar el Tipo de Prenda: Identificar las características de las diferentes prendas (tipo de tejido, grosor, etc.) para ajustar el ciclo de secado en consecuencia.

Ajustar el Ciclo de Secado: Modificar parámetros como la temperatura y la duración del ciclo según las necesidades específicas de la ropa.

Optimizar el Desempeño: Asegurar que cada ciclo de secado sea eficiente en términos de tiempo y consumo energético, manteniendo la calidad de las prendas.

##### Implementación de la Regresión Logística

Para alcanzar estos objetivos, implementaremos un modelo de regresión logística, que es adecuado para la clasificación binaria y multiclase. Este modelo ayudará a identificar el tipo de ciclo que la secadora debe utilizar basándose en las características de las prendas. Los pasos clave en la implementación son:

Recopilación de Datos: Utilizaremos conjuntos de datos reales proporcionados por la empresa MABE. Estos datos incluyen información detallada sobre diferentes tipos de prendas y los ciclos de secado asociados, permitiendo al modelo aprender las relaciones entre las características de la ropa y los ciclos de secado óptimos.

Preparación de Datos: Los datos se procesarán para asegurar que estén limpios y bien estructurados. Esto incluye la normalización de variables y el manejo de datos faltantes o anomalías.

Entrenamiento del Modelo: Se entrenará la regresión logística utilizando los datos recopilados. El modelo aprenderá a clasificar las prendas en función de sus características y predecir el ciclo de secado más adecuado.

Evaluación del Modelo: Se evaluará la precisión del modelo utilizando métricas de rendimiento como la exactitud, la precisión y la recuperación. Esto ayudará a ajustar los parámetros y mejorar la capacidad del modelo para generalizar a nuevas muestras.

##### Beneficios y Futuro

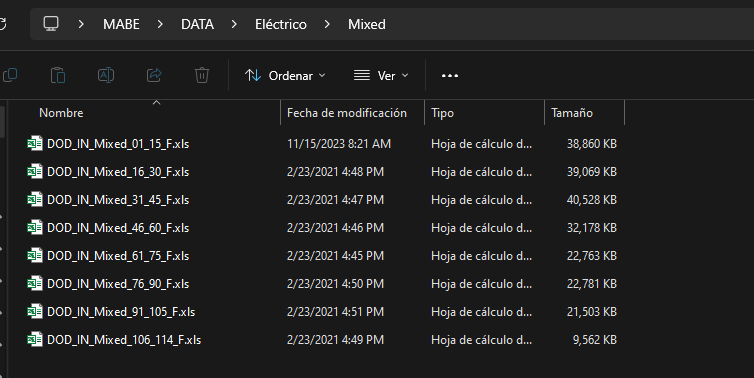
La implementación de la regresión logística en la secadora no solo promete mejorar la eficiencia del secado y la protección de las prendas, sino que también abre la puerta a futuras mejoras. Con el tiempo, el modelo puede refinarse y ajustarse utilizando datos adicionales y técnicas avanzadas de Machine Learning, como el aprendizaje profundo, para mejorar aún más la precisión y la adaptabilidad de la secadora.

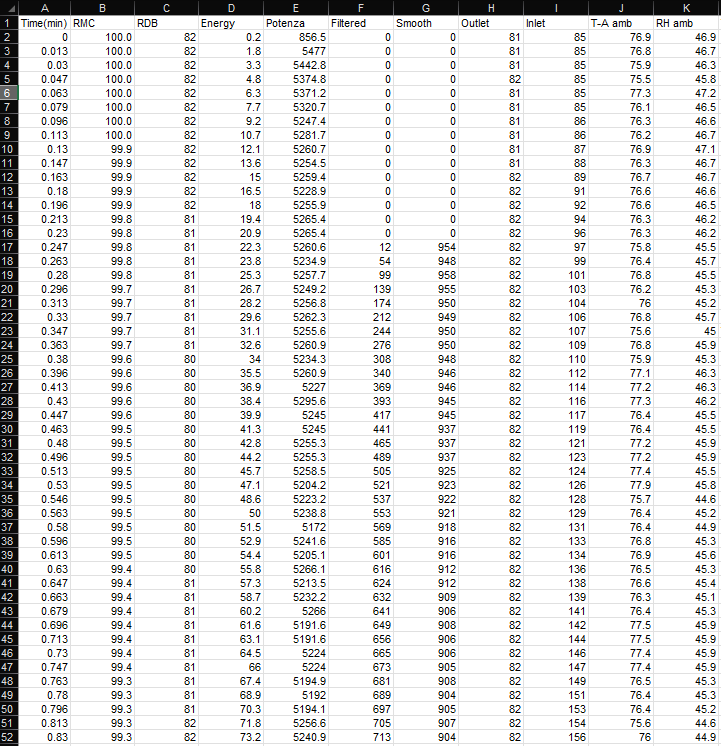
# ETL

ETL, Extract Transform Load por sus siglas en ingles, en esta sección explico como pase de varios archivos de Excel donde se registraban los resultados de diversos sensores instalados en una secadora, ah un Data Set que podemos usar para entrenar nuestra red logística.

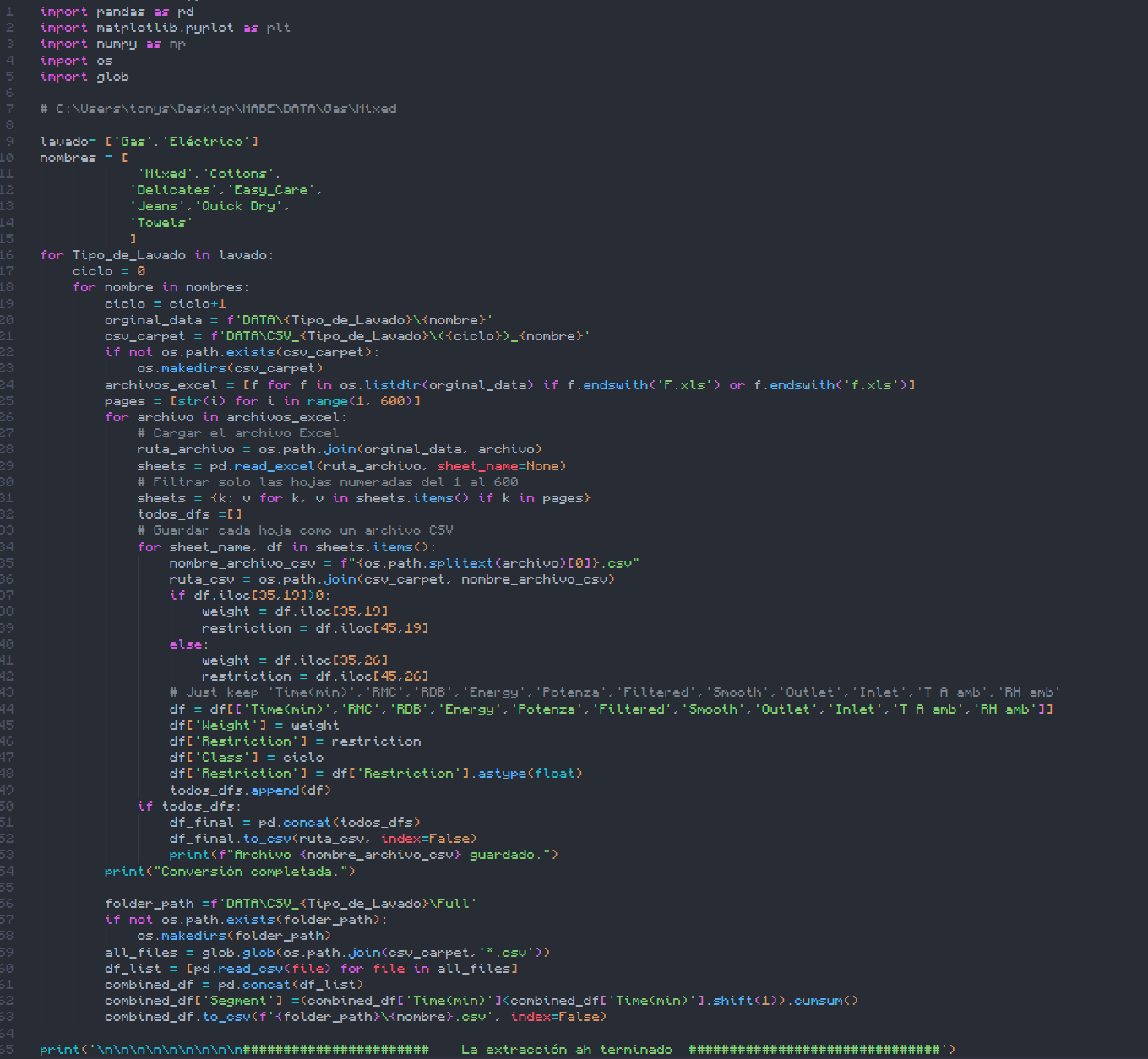
## Leyendo el experimento

Como mencioné el raw data me fue entregado en un formato tosco y complicado de leer, Excel tiene sus limitaciones a la hora de manejar grandes volúmenes de datos, sobre todo cuando se hace de manera incorrecta, es por esto que me dieron varios “Libros” de datos, y cada página de estos correspondía a un experimento o prueba realizada con la secadora.

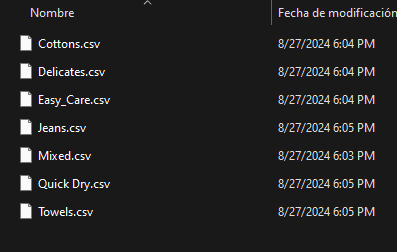




En total serian alrededor de 2721480 filas, por lo que primero decidí juntar todos estos libros en un solo data frame de Pandas, para poder agilizar cualquier operación que decidiera aplicar a esta información.



Este código recopila todos los libros, busca las páginas donde haya información de los experimentos, lee la ficha de cada experimento y recupera la masa de ropa y la restricción del flujo de aire del experimento. Después recupera las columnas que contiene la información del experimento, agrega ‘Clase’ y ‘Segmento’ que nos permite identificar a que ciclo y prueba de secado pertenece y concatena todo esto a un archivo CSV. Este proceso nos otorga una serie de CSV que podemos analizar de forma más cómoda.



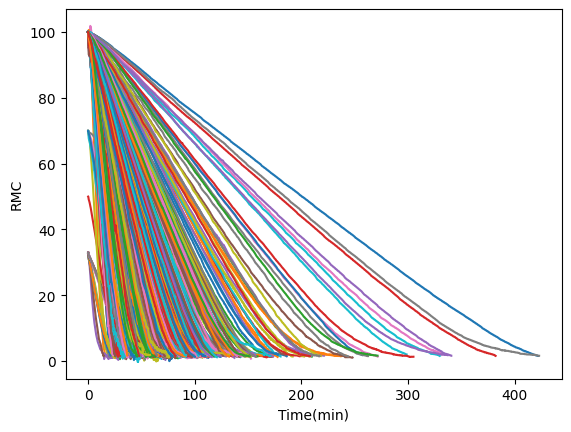
Ahora con los datos extraídos tocara decidir que variables son importantes o no para nuestra regresión logística.

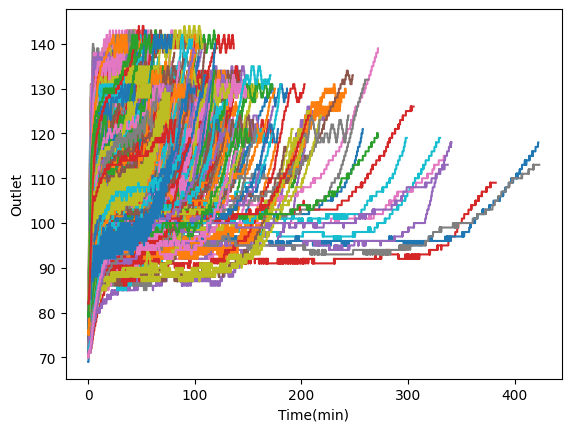
## Limpiando el Data Set

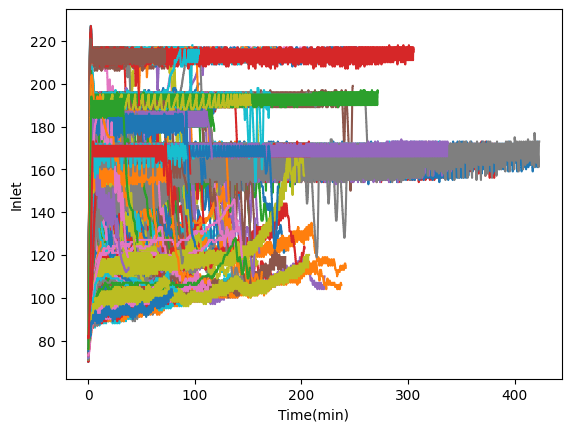
Para poder filtrar mas nuestro Data Set, hay que entender que representa cada variable de este. De esta manera podemos descartar que vale la pena mantener, que puede afectar y que no puede afectar el comportamiento de la secadora con cierto tipo de ropa.

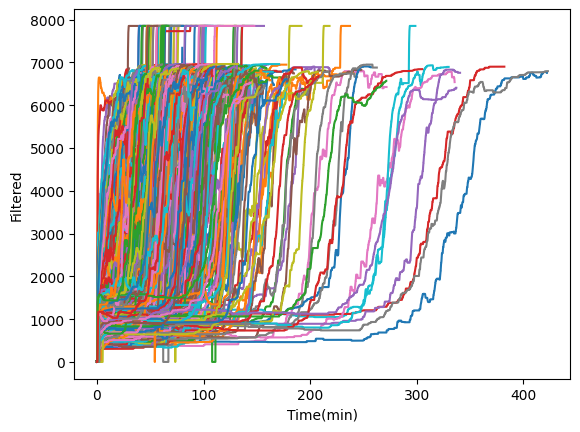
|  |  |
| --- | --- |
| Feature | Explicación |
| Time(min) | Se refiere al tiempo en minutos. |
| RMC | La humedad de la ropa, cuando esta llegue a un rango entre 4 y 2 % podemos considerar la ropa como seca. |
| RDB | Una segunda medición del valor de Outlet. |
| Energy | La energía en Volts durante la prueba. |
| Potenza | La potencia en Watts durante la prueba. |
| Filtered | Una señal filtrada. |
| Smooth | El suavizado de Filtered. |
| Outlet | La temperatura en F° del flujo de aire saliendo de la secadora. |
| Inlet | La temperatura producida por una resistencia eléctrica para secar la ropa. |
| T-A amb | Temperatura ambiental. |
| RH amb | Humedad ambiental. |
| Weight | La masa dentro de la secadora. |
| Restricction | La restricción del flujo de aire que escapa la secadora. |

Tal vez valdría la pena revisar el comportamiento de algunos de estos valores para saber que es lo que realmente sucede con ellos durante el ciclo de secado.









Aunque un poco saturadas estas graficas nos ayudan a comprender de forma general cómo se comportan estas 4 variables. Del RMC nos queda claro que siempre decae con el tiempo, lo que explica el comportamiento de Outlet, mientras menos humedad hay en el ambiente el aire a la salida de la secadora debería de tener una temperatura más alta. Por otro lado, el comportamiento de Inlet en un inicio parece independiente, puesto que trata de llegar a un setpoint dependiendo el ciclo y viéndose afectada por la masa de ropa, algo que podríamos tomar es su comportamiento una vez llega a su setpoint, ya que los picos se comportan de forma distinta. Por último, tenemos Filtered cuyo parece ser contrario al del RMC, lo que nos indicaría que este aumenta conforme menos humedad se encuentra en la secadora.

Después de revisar el comportamiento de cada variable, nos podemos dar una idea de que variables nos pueden ayudar a determinar la clase o el tipo de ropa dentro de la secadora. Por lo que conservaremos ‘RMC’,’Inlet’,’Outlet’,’Time(min)’ y ’Filtered’. Aun ocuparemos Features como ‘Weight’ o la ‘Restricction’ para enviar nuestras muestras en rangos específicos. En cuanto a la variables como la ‘T-A Amb’ , ‘RH Amb’ las descartamos por la poca diferencia que hay entre las muestras y los experimentos de modo que solo sesgarían al modelo al repetirse o tendrían un peso muy bajo. ‘Energy’ y ‘Potenza’ las descartamos por recomendación de MABE, nos comentaron que los sensores encargados de medir estas variables realmente no median algo significativo (‘la potencia del motor o de algún sistema del cual podamos inferir la clase de ropa).

Ah simple vista no parece que debamos quitar o arreglar alguna parte del Data Set, aunque tuvimos que intervenir en varios Excel para encontrar valores como la restricción o la masa del experimento.

Ahora toca decidir con que conjunto de datos probar nuestra regresión logística. Seria una buena idea encontrar datos que estadísticamente fueran diferentes, por lo que seria buena idea compara las descripciones estadísticas de nuestros experimentos.

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente

Gráfico, Gráfico de barras

Descripción generada automáticamente

Gráfico, Gráfico de barras

Descripción generada automáticamente

Gráfico, Gráfico de barras

Descripción generada automáticamente

De estas graficas podemos darnos cuenta de que las clases 5 y 7 son muy parecidas, tal vez convendría juntarlas en una sola para poder aumentar el tamaño de nuestro data set. La clase 6 y 2 difieren considerablemente de las demás clases. A continuación una recopilación de nuestros datos:

Interfaz de usuario gráfica

Descripción generada automáticamente

Interfaz de usuario gráfica

Descripción generada automáticamente

Si revisamos la cantidad de datos en cada Data Set nos podemos dar cuenta que podría resultar beneficioso usar la clase 4 y la clase 5, Jeans y Mixed, de esta forma nuestro modelo no se sesgaría y no tendríamos que recortar nuestro Data Set de forma tan dramática. Ahora solo debemos de transformar esta data.

Texto

Descripción generada automáticamente

Este codigo concatena ambos Data Frames, en los rangos que nosotros decidamos para la masa, o la restricción, aunque probando nos dimos cuenta que conviene más discriminar el peso antes que la restricción, puesto que este afecta más al ciclo de secado. Debemos darle una clase binaria a nuestro Data Set. Por ultimo agregamos algunas transformaciones a nuestros datos para que estos tengan una diferencia mayor entre Inlet y Outlet.

Pantalla de computadora con letras

Descripción generada automáticamente con confianza baja

Este seria el numero de filas para entrenar a nuestro modelo de regresión logística.

# Regresión Logística

Considerando que nuestro problema es un clasificador, conviene usar una regresión logística, ya que su función de activación sigmoide nos permite saber si algo pertenece a una de dos clases. En nuestro caso nos mostrara si pertenece a Mixed o a Jeans.

Ahora explicare como funciona nuestra modelo de Machine Learning.

### Cargar y separar los datos

A la hora de cargar nuestros datos nos conviene revolverlos para que el orden de estos no sesgue de alguna forma al modelo, separarlos en una proporción de 80:20 para entrenar y para probar, ‘Train’ y ‘Test’. También vale la pena escalar los datos desde este momento.

### Entrenamiento

El siguiente proceso lo repetiremos para cada ‘Epoch’.

#### Cálculo de nuevos parámetros

Primero, empezamos con una función hipótesis establecida de forma aleatoria. Se evalúan las predicciones hechas por esta función con los valores reales. Sacamos el gradiente para cada ‘Feature’, en nuestro caso lo hacemos de forma matricial con la función ‘.dot’ de numpy y las dividimos por la cantidad de muestras. Para obtener los nuevos parámetros multiplicamos este los respectivos gradientes por nuestro ‘learning rate’. Para sacar nuestro nuevo bias multiplicamos el error por nuestro learning rate.

#### Función de Costo

Ocupamos la entropía cruzada como función de costo, o perdida, esta función evalúa nuestra predicción de forma similar al error cuadrado que común mente ocupamos para regresiones lineales. De la misma manera buscamos llegar al punto mas bajo de nuestra función. Esta nos indicara que tan bien entrenado esta nuestro modelo, aunque no es una métrica tan útil para saber que tan correcto es nuestro modelo.

#### Validación por Accuracy

Este paso puede ser útil para saber que tan bien esta entrenado nuestro modelo con los datos de entrenamiento, y nos permite detectar fenómenos como el ‘Overfitting’. Evaluamos la predicción una sección especifica de nuestro Data Set y mostramos que porcentaje de aciertos tuvo.

Todo esto lo repetiremos por cada época hasta que lleguemos a un punto mínimo, se nos terminen las épocas o nos estanquemos con el error. Este ultimo se puede dar por muchos motivos, principalmente culpa del Data Set o del contexto del problema.

### Test

Por último, usamos la sección del Data Set que habíamos separado previamente para probar que tan bueno en nuestro modelo. Ocupando los parámetros y bias que nos haya entregado al final del entrenamiento y de la misma forma que para validar calculamos el ‘Accuracy’ de nuestro modelo.

Texto

Descripción generada automáticamente

Texto

Descripción generada automáticamente

# Interpretación y Resultados

Una vez hayamos corrido nuestro modelo deberíamos de terminar con el error general del modelo y que tan acertado es nuestro modelo, en nuestro caso, redondeamos la salida de nuestro modelo para evitar valores intermedios, en caso de aplicar el modelo esto también debería de hacerse. De esta forma obtenemos un error de 60.8% y un acertamiento del 70% lo cual es algo muy positivo.

Gráfico

Descripción generada automáticamente

Texto

Descripción generada automáticamente

Puede que agregando más transformaciones a nuestros ‘Features’ y agregando alguno que otro parámetro podamos mejorar estos números.

##### Referencias

1. Referencia [1]
2. Referencia [2]