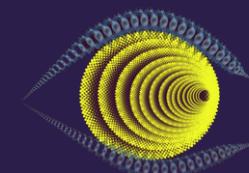
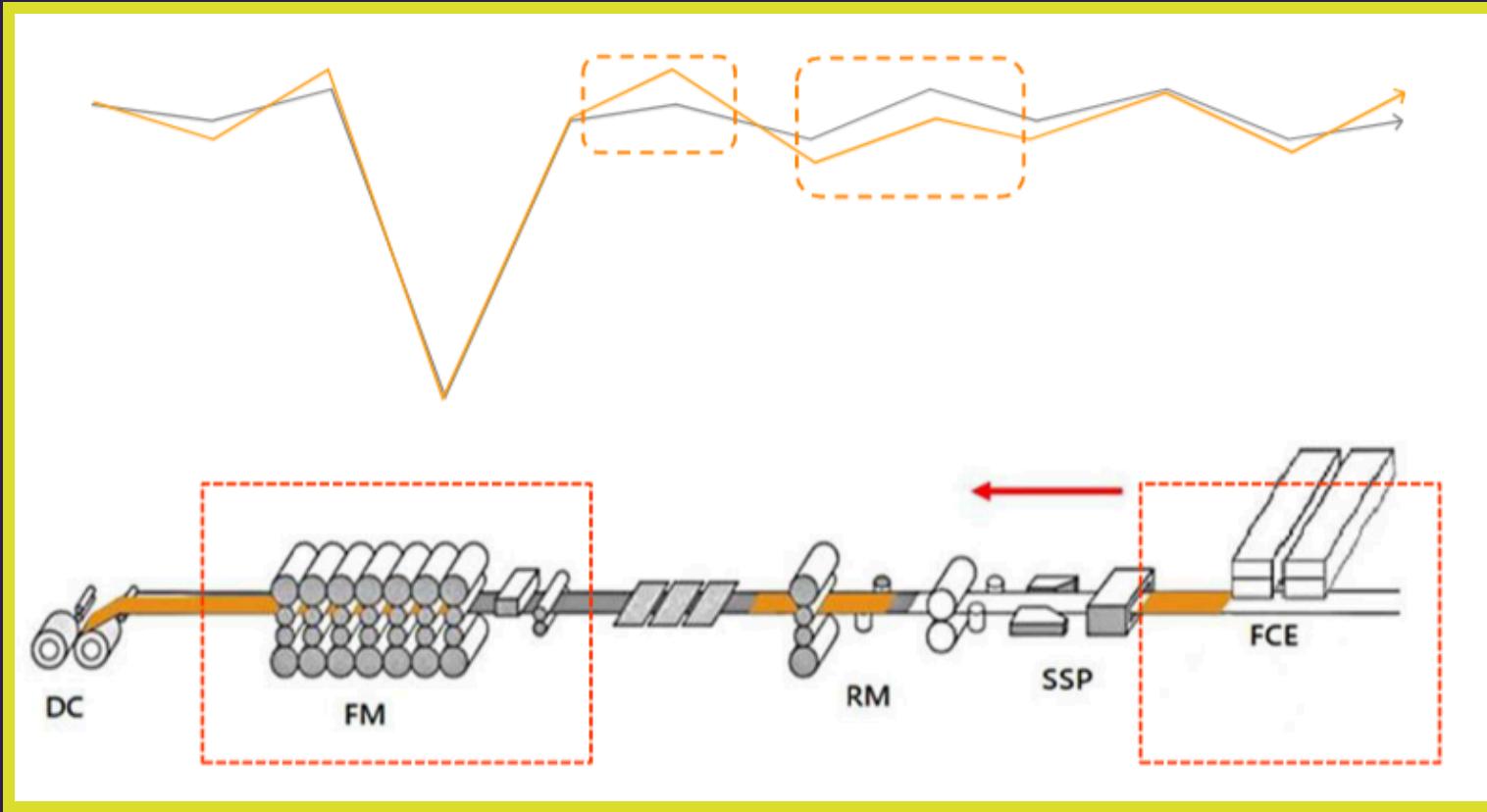


CHAMELEON

PRIMERA EVIDENCIA

2024

Cuellos de botella en el proceso de laminación del molino caliente 4

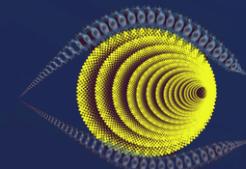


CHAMELEON

Unión de Bases

EXPORTACION_PACING_2024 + SGL_DEMORAS_CALIENTE 4 PES.

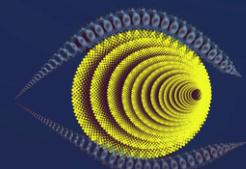
FM Travel Pacing Time Predicted	FM Proces Pacing Time Predicted	SSP use	SSP draft	Slab Width Nominal	Slab Thickness Nominal	Slab Length Nominal	Slab Weight Nominal	Demora	Interrupciones
37.0	75.0	1.0	83.32568	1570	230	7212	20.532	0	0
37.0	72.0	1.0	83.60986	1570	230	7211	20.529	0	0
35.0	83.0	1.0	59.74377	1570	220	8650	23.195	0	0
34.0	83.0	0.0	0.00000	1570	230	9510	26.937	0	0
34.0	52.0	0.0	0.00000	1570	230	8345	23.758	0	0



CHAMELEON

Limpieza

- *Eliminación columnas*
- *Tratamiento de NaN*
- *Tratamiento de 0's*
- *Eliminamos duplicados*
- *Agrupamos por familia en steel_type*
- *Convertimos la columna Bottleneck a numérica con mapeo*
- *Transformar a valores positivos los registros de Delay Time y F7-F1 Time*
- *Gráficos de distribución y boxplots*
- *Tratamiento de Outliers*



CHAMELEON

Limpieza

Eliminamos columnas:

- Fechas
- Predicted
- Yellow
- Nominal
- slab_id
- Comment y Subconcept

Tratamiento de NaN:

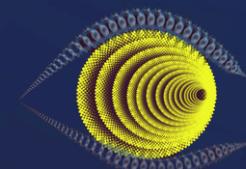
- Eliminar columnas con más del 30% de NaN
- Imputamos en las columnas numéricas con la mediana
- Borramos registros donde Commander Slab ID tenía NaN (9)

Tratamiento de 0's:

- Imputamos con la mediana en ciertas columnas numéricas
- Imputar los ceros en la columna Taper_Meas por la media

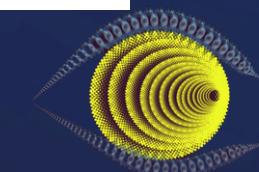
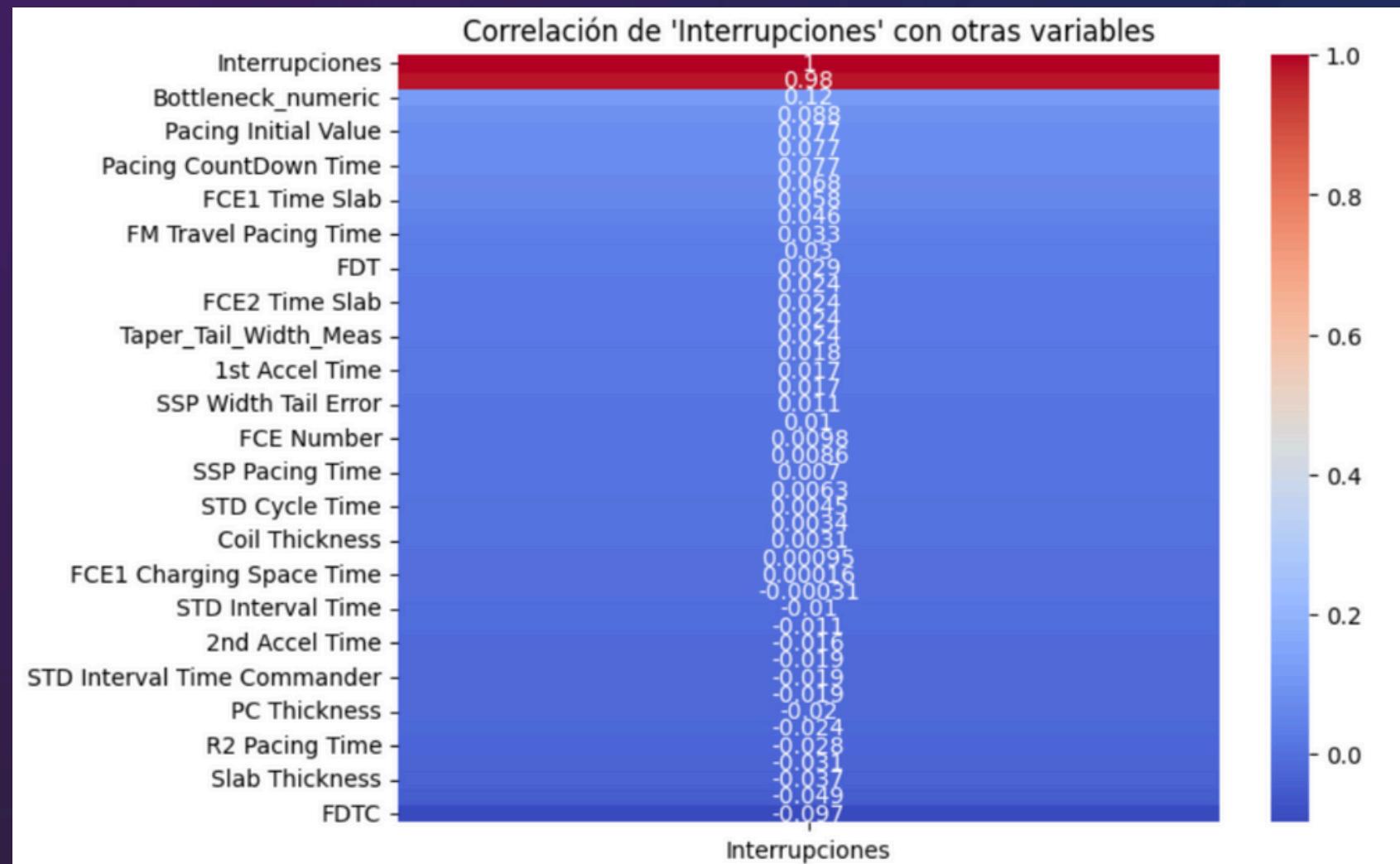
Tratamiento de outliers :

- Con IQR
- Reemplazar los outliers con NaN
- Rellenar los NaN resultantes con la mediana

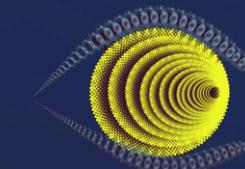
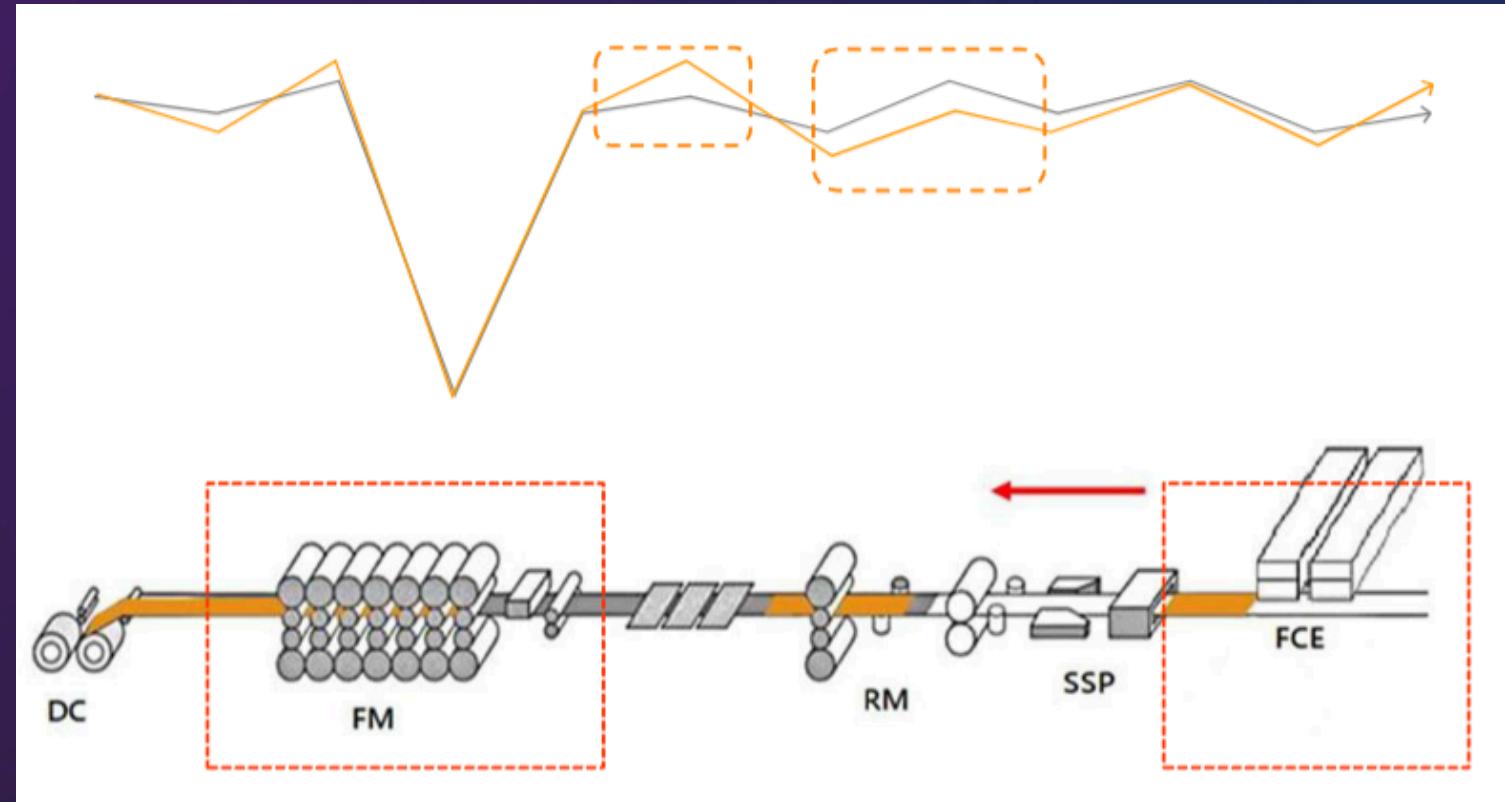


CHAMELEON

Matriz de correlación con Interrupciones



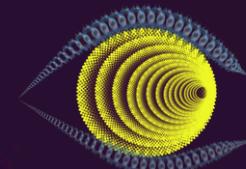
CHAMELEON



CHAMELEON

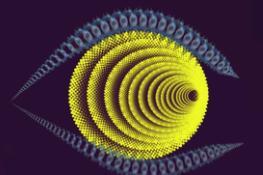
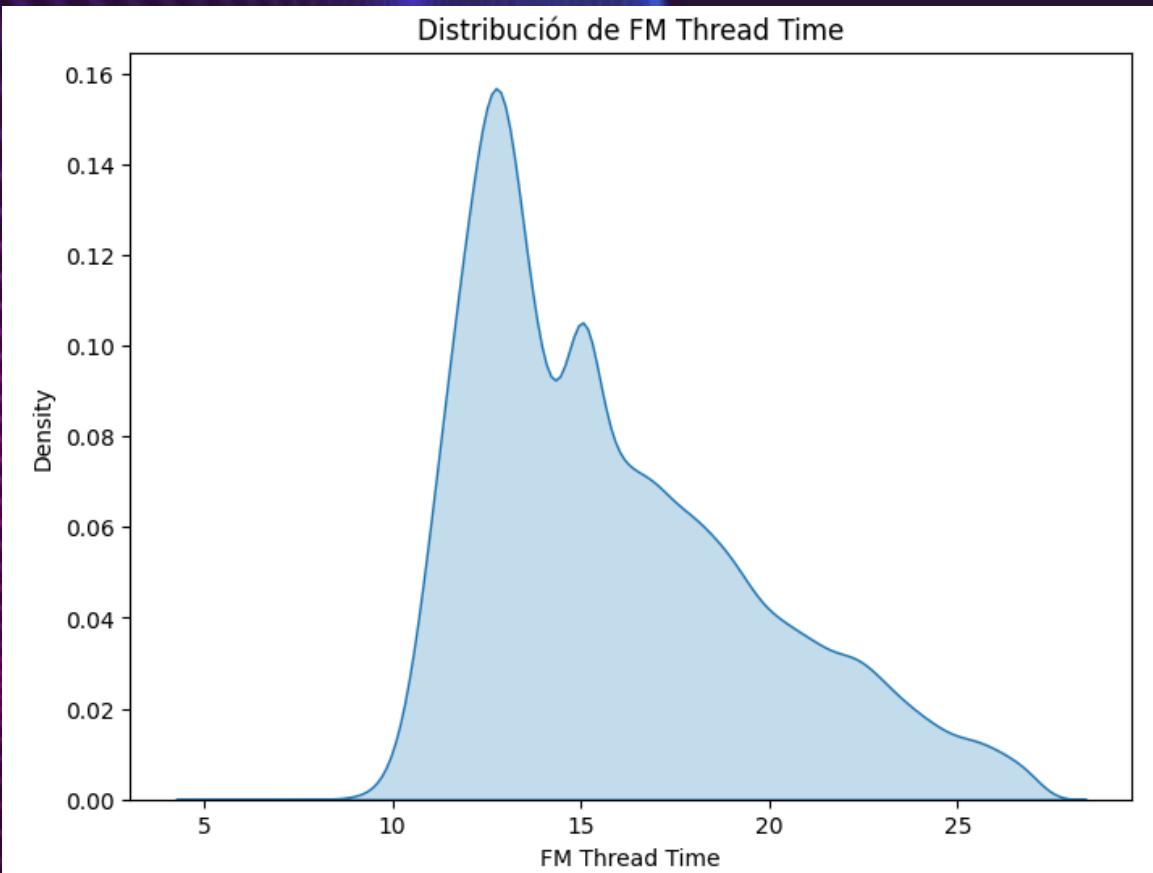
Correlaciones absolutas con FM Thread Time

	FM Thread Time
FM Thread Time	1.000000
FM Thread Speed	0.876033
FM Final Speed	0.773652
1st Accel Time	0.704730
Coil Thickness	0.692724
FM Travel Pacing Time	0.575000
R2 Pacing Time	0.468276
STD Interval Time	0.467072
steel_type	0.443852
2nd Accel Time	0.436825
Slab Width	0.368901
Taper_Head_Width_Meas	0.362130
Taper_Tail_Width_Meas	0.362009
FDT	0.350008
STD Cycle Time	0.298635
SSP use	0.271344
Slab Weight	0.214291
Slab Length	0.212053



CHAMELEON

Distribución de FM Thread Time



CHAMELEON

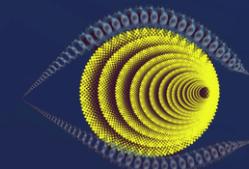
Regresión Lineal Múltiple

```
X = data[['FM Thread Speed', 'FM Final Speed', '1st Accel Time', 'FM Travel Pacing Time', 'R2 Pacing Time', 'steel_type', 'Slab Width',  
'Taper_Head_Width_Meas', 'Taper_Tail_Width_Meas', 'Slab Weight', 'Slab Length', 'Slab Thickness']]  
y = data['FM Thread Time']
```

```
Mean Squared Error (MSE): 2.6713390844233817  
R-squared (R2): 0.8145730523951208
```

Coeficientes del modelo:

	Coeficiente
FM Thread Speed	-1.315422
FM Final Speed	0.242559
1st Accel Time	0.046614
FM Travel Pacing Time	0.282664
R2 Pacing Time	-0.003016
steel_type	0.096135
Slab Width	-0.001023
Taper_Head_Width_Meas	0.001696
Taper_Tail_Width_Meas	-0.002070
Slab Weight	0.112008
Slab Length	-0.000117
Slab Thickness	-0.017385



CHAMELEON

Regresión Logística

$y = 1$ si 'FM Thread Time' > media, 0 si no

```
X = data[['FM Thread Speed', 'FM Final Speed', '1st Accel Time', 'FM Travel Pacing Time', 'R2 Pacing Time', 'steel_type',  
'Slab Width', 'Taper_Head_Width_Meas', 'Taper_Tail_Width_Meas', 'Slab Weight', 'Slab Length', 'Slab Thickness']]  
y = (data['FM Thread Time'] > data['FM Thread Time'].mean()).astype(int)
```

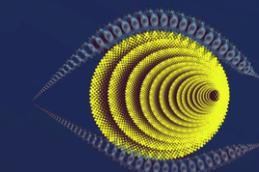
```
Accuracy: 0.90529198692464  
Matriz de Confusión:  
[[6219 495]  
 [ 577 4028]]  
Reporte de Clasificación:  
 precision recall f1-score support  
  
 0 0.92 0.93 0.92 6714  
 1 0.89 0.87 0.88 4605  
  
 accuracy 0.91 11319  
 macro avg 0.90 0.90 11319  
 weighted avg 0.91 0.91 11319  
  
Coeficientes del modelo:  
 Coeficiente  
FM Thread Speed -0.518857  
FM Final Speed -0.124506  
1st Accel Time 0.140079  
FM Travel Pacing Time 0.314225  
R2 Pacing Time -0.052414  
steel_type -0.321845  
Slab Width -0.001563  
Taper_Head_Width_Meas 0.002114  
Taper_Tail_Width_Meas -0.002104  
Slab Weight 0.109714  
Slab Length -0.000050  
Slab Thickness -0.016635
```

Modelo KNN

$y = 1$ si 'FM Thread Time' > media, 0 si no

```
X = data[['FM Thread Speed', 'FM Final Speed', '1st Accel Time', 'FM Travel Pacing Time', 'R2 Pacing Time', 'steel_type',  
'Slab Width', 'Taper_Head_Width_Meas', 'Taper_Tail_Width_Meas', 'Slab Weight', 'Slab Length', 'Slab Thickness']]  
y = (data['FM Thread Time'] > data['FM Thread Time'].mean()).astype(int)
```

```
Resultados para K=3:  
[[6323 391]  
 [ 321 4284]]  
 precision recall f1-score support  
  
 0 0.95 0.94 0.95 6714  
 1 0.92 0.93 0.92 4605  
  
 accuracy 0.94 11319  
 macro avg 0.93 0.94 0.93 11319  
 weighted avg 0.94 0.94 0.94 11319  
  
Resultados para K=5:  
[[6332 382]  
 [ 293 4312]]  
 precision recall f1-score support  
  
 0 0.96 0.94 0.95 6714  
 1 0.92 0.94 0.93 4605  
  
 accuracy 0.94 11319  
 macro avg 0.94 0.94 0.94 11319  
 weighted avg 0.94 0.94 0.94 11319  
  
Resultados para K=10:  
[[6374 340]  
 [ 378 4227]]  
 precision recall f1-score support  
  
 0 0.94 0.95 0.95 6714  
 1 0.93 0.92 0.92 4605  
  
 accuracy 0.94 11319  
 macro avg 0.93 0.93 0.93 11319  
 weighted avg 0.94 0.94 0.94 11319
```



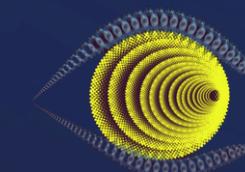
CHAMELEON

Modelo Random Forest

$y = 1$ si 'FM Thread Time' > media, 0 si no

```
X = data[['FM Thread Speed', 'FM Final Speed', '1st Accel Time', 'FM Travel Pacing Time', 'R2 Pacing Time', 'steel_type',  
'Slab Width', 'Taper_Head_Width_Meas', 'Taper_Tail_Width_Meas', 'Slab Weight', 'Slab Length', 'Slab Thickness']]  
y = (data['FM Thread Time'] > data['FM Thread Time'].mean()).astype(int)
```

```
Accuracy: 0.9529110345436876  
Matriz de Confusión:  
[[6417 297]  
 [ 236 4369]]  
Reporte de Clasificación:  
precision recall f1-score support  
  
0 0.96 0.96 0.96 6714  
1 0.94 0.95 0.94 4605  
  
accuracy 0.95 0.95 0.95 11319  
macro avg 0.95 0.95 0.95 11319  
weighted avg 0.95 0.95 0.95 11319  
  
Importancia de las características:  
Importancia  
FM Thread Speed 0.318961  
FM Final Speed 0.245184  
1st Accel Time 0.174903  
FM Travel Pacing Time 0.061374  
steel_type 0.043197  
R2 Pacing Time 0.031107  
Slab Width 0.023907  
Taper_Tail_Width_Meas 0.022665  
Slab Length 0.022489  
Slab Weight 0.021549  
Taper_Head_Width_Meas 0.018854  
Slab Thickness 0.015811
```

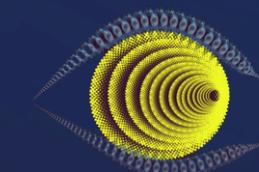
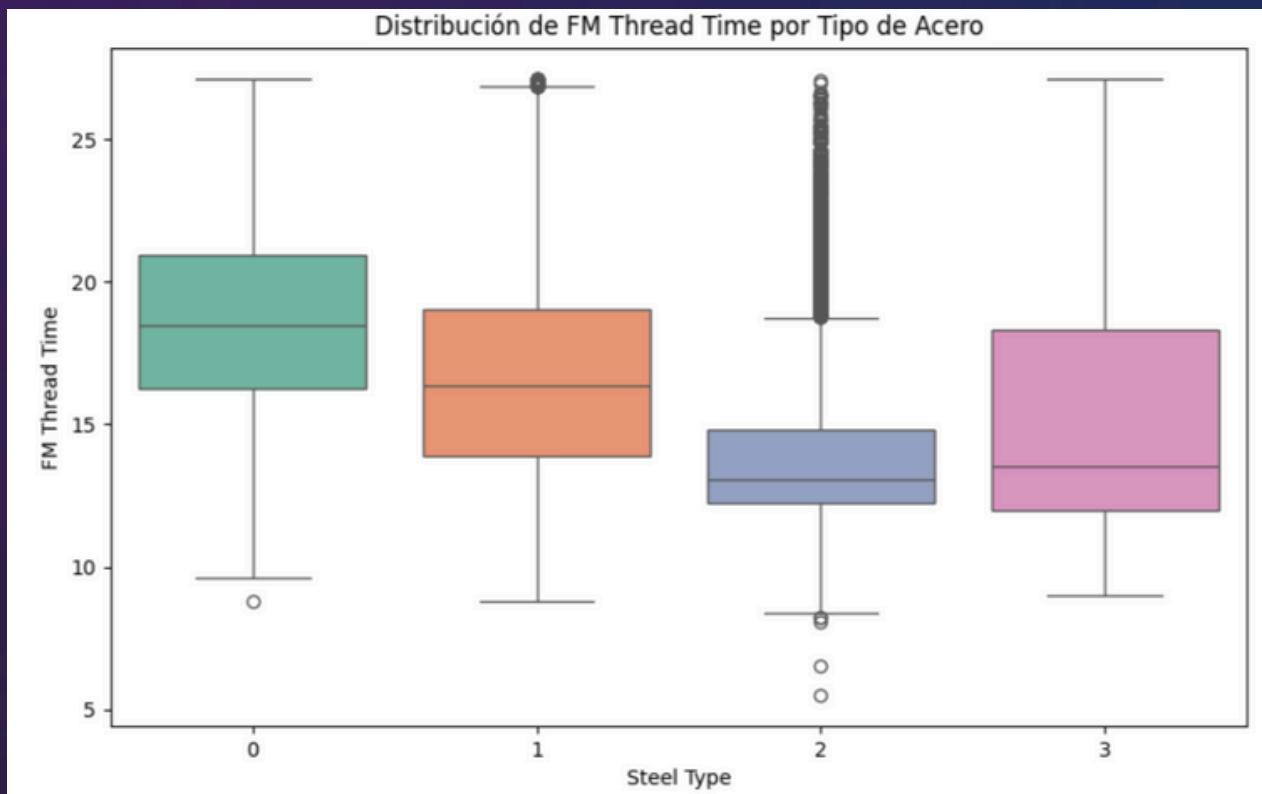


CHAMELEON

ANOVA para la variable 'FM Thread Time' según 'steel_type'

Resultados del ANOVA:

F-statistic: 7303.531137960513, p-value: 0.0



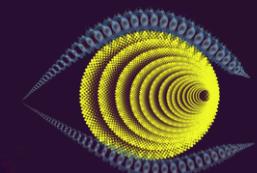
CHAMELEON

ANOVA entre modelos

Lineal Múltiple, Logístico y Random Forest

- Hay diferencias estadísticamente significativas entre las predicciones de los tres modelos.
- Los modelos están capturando diferentes aspectos de los datos.
- Si uno de los modelos tiene mejores métricas (precisión, R^2 , etc.), ANOVA simplemente confirma que hay un modelo que predice mejor que los otros.

```
ANOVA results: F_onewayResult(statistic=221704.26105657843, pvalue=0.0)
```



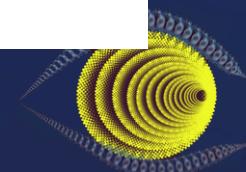
CHAMELEON

Intervalos de confianza del modelo de random forest

- Todas las importancias permutadas de las variables que calculamos están dentro de sus respectivos intervalos de confianza.
- Cuanto más estrecho sea el intervalo la importancia real de la variable es mayor.
- FM Thread Speed y FM Travel Pacing Time son las variables que más influyen en los cuellos de botella porque su intervalo es más estrecho.

Importancia permutada de las características con intervalos de confianza:

FM Thread Speed: 0.2191 (95% CI: 0.2119, 0.2262)	0.0143
FM Travel Pacing Time: 0.0459 (95% CI: 0.0425, 0.0493)	0.0068
FM Final Speed: 0.0110 (95% CI: 0.0093, 0.0127)	0.0034
1st Accel Time: 0.0090 (95% CI: 0.0074, 0.0106)	0.0032
steel_type: 0.0057 (95% CI: 0.0039, 0.0075)	0.0036
R2 Pacing Time: 0.0051 (95% CI: 0.0038, 0.0064)	0.0026
Slab Thickness: 0.0047 (95% CI: 0.0035, 0.0059)	0.0024
Slab Weight: 0.0033 (95% CI: 0.0021, 0.0046)	0.0025
Slab Length: 0.0031 (95% CI: 0.0016, 0.0046)	0.003
Slab Width: 0.0026 (95% CI: 0.0016, 0.0036)	0.002
Taper_Tail_Width_Meas: 0.0026 (95% CI: 0.0017, 0.0034)	0.0017
Taper_Head_Width_Meas: 0.0015 (95% CI: 0.0003, 0.0026)	0.0023

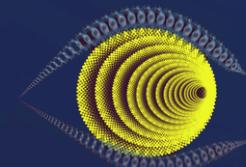


Intervalos de confianza del modelo de regresión logística

- Si el intervalo de confianza no contiene el valor 0, la variable es estadísticamente significativa.
- Si el intervalo de confianza incluye el valor 0, la variable no es estadísticamente significativa.
- Todas las variables tienen intervalos de confianza que no incluyen el valor 0, lo que significa que son estadísticamente significativas.

Intervalos de confianza para los coeficientes del modelo de regresión logística:

```
FM Thread Speed: (-0.5301, -0.5076)
FM Final Speed: (-0.1309, -0.1182)
1st Accel Time: (0.1387, 0.1415)
FM Travel Pacing Time: (0.3100, 0.3184)
R2 Pacing Time: (-0.0551, -0.0498)
steel_type: (-0.3329, -0.3108)
Slab Width: (-0.0018, -0.0013)
Taper_Head_Width_Meas: (0.0009, 0.0033)
Taper_Tail_Width_Meas: (-0.0033, -0.0009)
Slab Weight: (0.1032, 0.1163)
Slab Length: (-0.0001, -0.0000)
Slab Thickness: (-0.0175, -0.0158)
```



Evaluación en conjuntos de entrenamiento y prueba

Regresión Lineal Múltiple

Conjunto de entrenamiento:

Mean Squared Error (MSE): 2.563908576877576

R-squared (R²): 0.8241628857280849

Conjunto de prueba:

Mean Squared Error (MSE): 2.6713390844233817

R-squared (R²): 0.8145730523951208

Coeficientes del modelo:

	Coeficiente
FM Thread Speed	-1.315422
FM Final Speed	0.242559
1st Accel Time	0.046614
FM Travel Pacing Time	0.282664
R2 Pacing Time	-0.003016
steel_type	0.096135
Slab Width	-0.001023
Taper_Head_Width_Meas	0.001696
Taper_Tail_Width_Meas	-0.002070
Slab Weight	0.112008
Slab Length	-0.000117
Slab Thickness	-0.017385

Regresión Logística

Métricas de entrenamiento:

Accuracy: 0.9147009453131902

Matriz de Confusión:

```
[[24669 1750]
 [ 2112 16745]]
```

Reporte de Clasificación:

	precision	recall	f1-score	support
0	0.92	0.93	0.93	26419
1	0.91	0.89	0.90	18857
accuracy			0.91	45276
macro avg	0.91	0.91	0.91	45276
weighted avg	0.91	0.91	0.91	45276

Métricas de prueba:

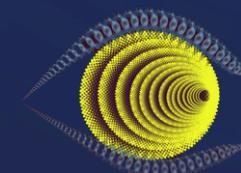
Accuracy: 0.90529198692464

Matriz de Confusión:

```
[[6219 495]
 [ 577 4028]]
```

Reporte de Clasificación:

	precision	recall	f1-score	support
0	0.92	0.93	0.92	6714
1	0.89	0.87	0.88	4605
accuracy			0.91	11319
macro avg	0.90	0.90	0.90	11319
weighted avg	0.91	0.91	0.91	11319

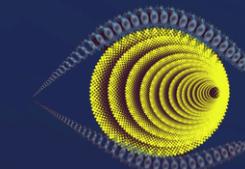


CHAMELEON

Conclusiones

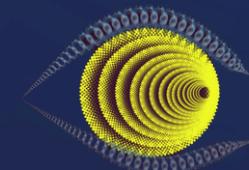
- El modelo con mejor desempeño fue un modelo de clasificación, y fue el random forest con una precisión de 95%, y un f1-score (media armónica de precisión y recall) de 0.96 para la clase 0 y de 0.91 para la clase 1
- Las variables con mayor correlación con FM Thread Time fueron: FM Thread Speed, FM Final Speed y 1st Accel Time.
- FM Thread Speed y FM Travel Pacing Time son las variables que más influyen en los cuellos de botella porque su intervalo es más estrecho, pero el orden de steel_type y los patrones de medidas de los planchones son las variables que podrían controlar.

Importancia de las características:	Importancia
FM Thread Speed	0.318961
FM Final Speed	0.245184
1st Accel Time	0.174903
FM Travel Pacing Time	0.061374
<u>steel_type</u>	0.043197
R2 Pacing Time	0.031107
Slab Width	0.023907
Taper_Tail_Width_Meas	0.022665
Slab Length	0.022489
Slab Weight	0.021549
Taper_Head_Width_Meas	0.018854
Slab Thickness	0.015811



Siguientes Pasos

- Modelos K-Means y PCA.
- Dividir la BD para hacer un análisis por fases en el proceso del MC4.
- Simulación de escenarios para identificar patrones que influyan en los cuellos de botella.
- Análisis de fases por familias de acero.



CHAMELEON

CONTACTO

TELÉFONO

(123) 456-7890

EMAIL

chameleon@gmail.com

SOCIAL

