

EQUIPO 4 | Máquinas de chambeo





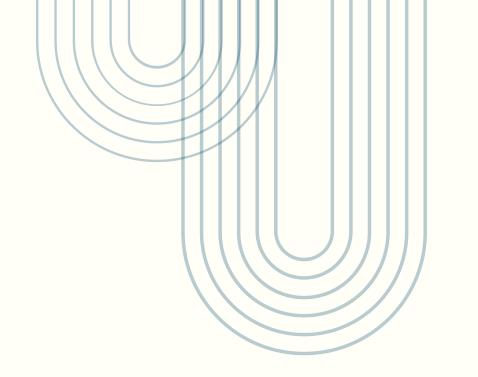


INICIO DE CÓDIGO

- Importamos las bibliotecas esenciales para el análisis
- Realizamos la carga de los datos por el archivo csv
- Realizamos la exploración de datos, y la búsqueda de valores nulos por columna

NULOS Y OUTLIERS

- Varias columnas tienen valores nulos, especialmente:
- "botón correcto", "tiempo de interacción", "número de interacción", "color presionado" y "auto push" con 762 valores nulos cada una.



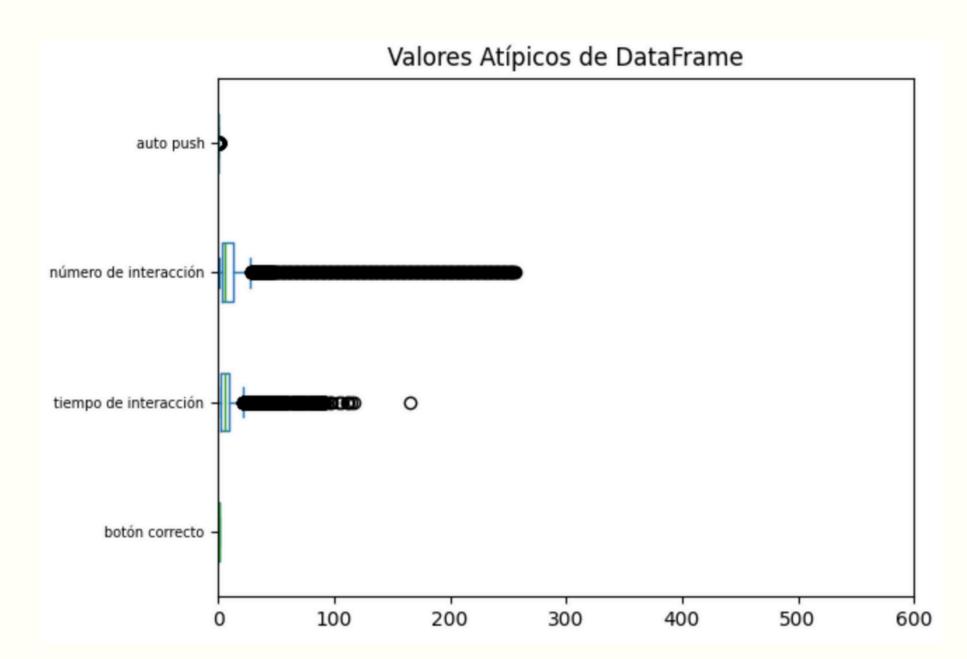


- Se **verifican** los tipos de **datos** de cada columna
- Separamos las variables por numéricas y cuantitativas:
 - Cuantitativas ('int64' & 'float64'): media
 - Cualitativas ('object'): "Sin dato"





Valores Atípicos del DataFrame



Valores Atípicos de DataFrame

Variables: auto push, número de interacción, tiempo de interacción y botón correcto

Método: rango intercuartílico (IQR)

Variables: tiempo de lección y tiempo de sesión (por la gran cantidad de registros "O")

Método: percentil 1 y 99

CONVERSIÓN DE VARIABLES

categóricas — numéricas

método Frecuencia

Administrador	
ALEIDA	3260
nicolas	440
LEONARDO	371
DENISSE	302
SERGIO ANGEL	243
CARLOS ENRIQUE	228
YAEL DAVID	224
AUSTIN	199
VALENTIN	163
erick	158
IKER BENJAMIN	128
KYTZIA	98
BENJAMIN	51
Name: Administra	dor, dtype: int64

Aleida	1
Nicolas	2
Leonardo	3
Denisse	4
Sergio Angel	5

ESTRATEGIA DE SELECCIÓN

de variables a evaluar

- Revisamos el heatmap para encontrar las variables más relevantes
- Escogimos las variables objetivo y variables independientes que tuvieran los coeficientes más altos dentro de los heatmaps
- Aplicamos diferentes modelos no lineales para evaluar si era posible aumentar los coeficientes
- Escogimos aquellos modelos que elevaron más los coeficientes
- Comparamos los resultados



	- Administrador	- Usuario	- mini juego	- color presionado	- dificultad	- Juego	- botón correcto	- tiempo de interacción	- número de interacción	- tiempo de sesión	- tiempo de lección
Administrador -	1.00	0.32	0.05	0.04	0.10	0.08	0.16	0.06	0.17	0.01	0.06
Usuario -	0.32	1.00	0.23	0.05	0.08	0.32	0.14	0.13	0.13	0.03	0.05
mini juego -	0.05	0.23	1.00	0.18	0.16	0.63	0.08	0.22	0.01	0.27	0.10
color presionado -	0.04	0.05	0.18	1.00	0.03	0.06	0.00	0.20	0.24	0.29	0.45
dificultad -	0.10	0.08	0.16	0.03	1.00	0.12	0.00	0.14	0.11	0.08	0.04
Juego -	0.08	0.32	0.63	0.06	0.12	1.00	0.05	0.21	0.06	0.02	0.09
botón correcto -	0.16	0.14	0.08	0.00	0.00	0.05	1.00	0.03	0.26	0.00	0.00
tiempo de interacción -	0.06	0.13	0.22	0.20	0.14	0.21	0.03	1.00	0.16	0.16	0.24
número de interacción -	0.17	0.13	0.01	0.24	0.11	0.06	0.26	0.16	1.00	0.19	0.29
tiempo de sesión -	0.01	0.03	0.27	0.29	0.08	0.02	0.00	0.16	0.19	1.00	0.16
tiempo de lección -	0.06	0.05	0.10	0.45	0.04	0.09	0.00	0.24	0.29	0.16	1.00

Heatmap General

Variable x	mini juego
Variable y	Juego
Modelo	Exponencial

COEFICIENTES

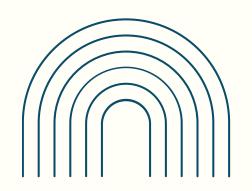
REGRESIÓN LINEAR SIMPLE

MODELO NO LINEAR

0.64



0.68



RESULTADOS POR USUARIO

Benjamín

	- mini juego	- color presionado	- dificultad	- botón correcto	- tiempo de interacción	- número de interacción	- tiempo de sesión	- tiempo de lección	1.0
mini juego -	1.00	0.26	0.08	0.02	0.10	0.21	0.87	0.03	
color presionado -	0.26	1.00	-0.33	-0.14	0.30	0.49	0.34	0.51	- 0.8
dificultad -	0.08	-0.33	1.00	0.04	0.04	0.12	0.30	0.01	- 0.6
botón correcto -	0.02	-0.14	0.04	1.00	-0.07	0.16	0.04	0.05	- 0.4
tiempo de interacción -	0.10	0.30	0.04	-0.07	1.00	0.33	0.08	0.12	- 0.2
número de interacción -	0.21	0.49	0.12	0.16	0.33	1.00	0.35	0.51	
tiempo de sesión -	0.87	0.34	0.30	0.04	0.08	0.35	1.00	0.09	- 0.0
tiempo de lección -	0.03	0.51	0.01	0.05	0.12	0.51	0.09	1.00	0.2

Variables	Función	Coef. lineal	Coef. no lineal
x = número de interacción y = tiempo de interacción	cuadrática	0.33	0.59
x = número de interacción y = botón correcto	exponencial	0.16	0.32
x = número de interacción y = color presionado	cuadrática	0.49	0.54
x = tiempo de sesión y = dificultad	cuadrática	0.30	0.34

Carlos Abel

	- mini juego	- color presionado	- dificultad	- botón correcto	- tiempo de interacción	- número de interacción	- tiempo de sesión	- tiempo de lección	- 1.0
mini juego -	1.00	0.27	-0.20	0.34	-0.26	0.03	0.12	-0.13	
color presionado -	0.27	1.00	0.29	0.11	-0.06	0.25	0.62	0.41	- 0.8
dificultad -	-0.20	0.29	1.00	-0.25	0.29	0.06	0.01	0.20	- 0.6
botón correcto -	0.34	0.11	-0.25	1.00	-0.53	-0.16	-0.04	-0.03	- 0.4
tiempo de interacción -	-0.26	-0.06	0.29	-0.53	1.00	0.50	0.20	0.14	- 0.2
número de interacción -	0.03	0.25	0.06	-0.16	0.50	1.00	0.76	0.50	- 0.0
tiempo de sesión -	0.12	0.62	0.01	-0.04	0.20	0.76	1.00	0.67	0.2
tiempo de lección -	-0.13	0.41	0.20	-0.03	0.14	0.50	0.67	1.00	0.4

Variables	Función	Coef. lineal	Coef. no lineal
x = número de interacción y = tiempo de interacción	cuadrática	0.50	0.61
x = color presionado y = botón correcto	exponencial	0.11	0.23
x = tiempo de sesión y = color presionado	cuadrática	0.62	0.65
x = color presionado y = dificultad	exponencial	0.29	0.36

Carlos Enrique

	- mini juego	- color presionado	- dificultad	oben[-	- botón correcto	- tiempo de interacción	- número de interacción	- tiempo de sesión	- tiempo de lección
mini juego -	1.00	0.23	-0.21	0.55	-0.16	0.35	0.22	0.45	0.09
color presionado -	0.23	1.00	0.03	-0.02	-0.39	0.69).68	0.43	0.55
dificultad -	-0.21	0.03	1.00	-0.37	-0.13	0.00	0.17	0.21	-0.04
Juego -	0.55	-0.02	-0.37	1.00	0.03	0.06	-0.09	-0.14	0.10
botón correcto -	-0.16	-0.39	-0.13	0.03	1.00	-0.50	-0.57	-0.34	-0.43
tiempo de interacción -	0.35	0.69	0.00	0.06	-0.50	1.00	0.80	0.53	0.67
número de interacción -	0.22	0.68	0.17	-0.09	-0.57	0.80	1.00	0.54	0.69
tiempo de sesión -	0.45	0.43	0.21	-0.14	-0.34	0.53	0.54	1.00	0.20
tiempo de lección -	0.09	0.55	-0.04	0.10	-0.43	0.67	0.69	0.20	1.00

- 0.2

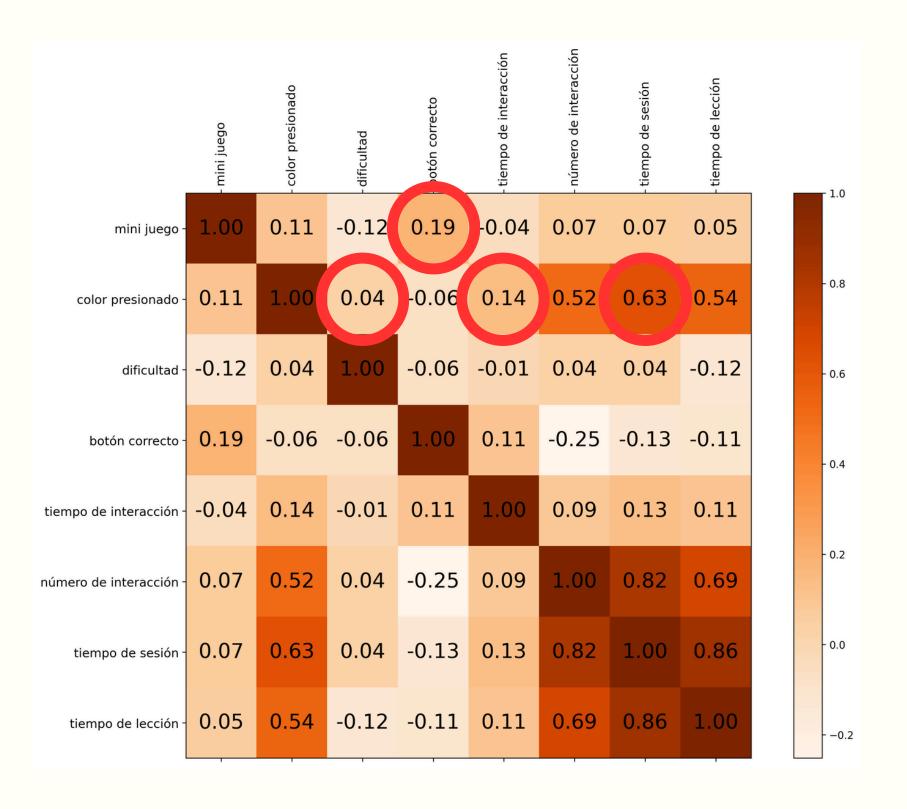
- 0.0

- -0.2

-0.4

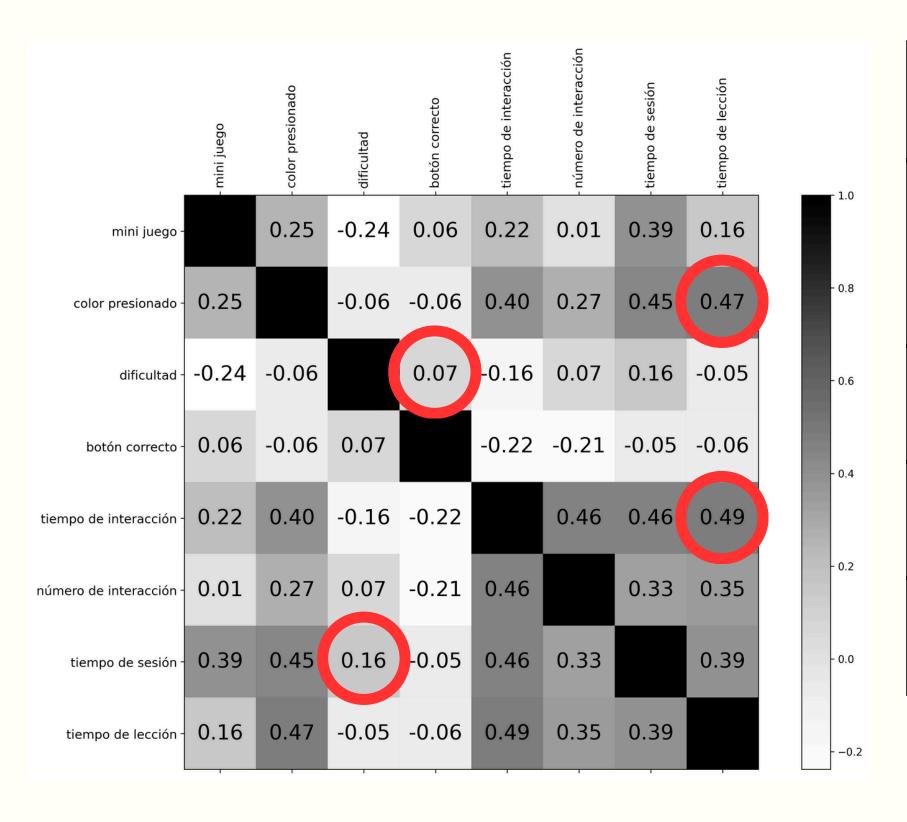
Variables	Función	Coef. lineal	Coef. no lineal
x = número de interacción y = tiempo de interacción	cuadrática	0.80	0.88
x = dificultad y = botón correcto	cuadrática	-0.13	0.14
x = tiempo de interacción y = color presionado	cuadrática	0.69	0.71
x = número de interacción y = dificultad	cuadrática	0.17	0.23

Concepción



Variables	Función	Coef. lineal	Coef. no lineal
x = color presionado y = tiempo de interacción	cuadrática	0.14	0.24
x = mini juego y = botón correcto	senoidal	0.19	0.23
x = tiempo de sesión y = color presionado	cuadrática	0.63	0.66
x = color presionado y = dificultad	cuadrática	0.04	0.08

Denisse



Variables	Función	Coef. lineal	Coef. no lineal
x = tiempo de lección y = tiempo de interacción	cuadrática	0.49	0.60
x = dificultad y = botón correcto	cuadrática	0.07	0.10
x = tiempo de lección y = color presionado	cuadrática	0.47	0.58
x = tiempo de sesión y = dificultad	cuadrática	0.16	0.18

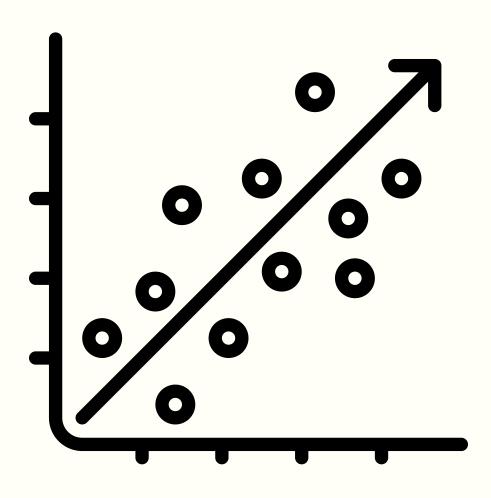
Usuario	Variable	Modelo no lienal	Coef. Lineal	Coef. No Lineal	Aumento
	x número de interacción y tiempo de interacción	cuadrática	0.33	0.59	79%
	x número de interacción y botón correcto	exponencial	0.16	0.32	100%
	x número de interacción y color presionado	cuadrática	0.49	0.54	10%
	x tiempo de sesión y dificultad	cuadrática	0.3	0.34	13%
	x número de interacción y tiempo de interacción	cuadrática	0.5	0.61	22%
	x color presionado y botón correcto	exponencial	0.11	0.23	109%
	x tiempo de sesión y color presionado	cuadrática	0.62	0.65	5%
	x color presionado y dificultad	exponencial	0.29	0.36	24%
	x número de interacción y tiempo de interacción	cuadrática	0.8	0.88	10%
	 x dificultad y botón correcto x tiempo de interacción y color presionado 	cuadrática	-0.13	0.14	114%
		cuadrática	0.69	0.71	3%
	x número de interacción y dificultad	cuadrática	0.17	0.23	35%
	x color presionado y tiempo de interacción	cuadrática	0.14	0.24	71%
	x mini juego y botón correcto	senoidal	0.19	0.23	21%
	x tiempo de sesión y color presionado	xua	0.63	0.66	5%
	x color presionado y dificultad	cuadrática	0.04	0.08	100%
	x tiempo de lección y tiempo de interacción	cuadrática	0.49	0.6	22%
	x dificultad y botón correcto	cuadrática	0.07	0.1	43%
	x tiempo de lección y color presionado	cuadrática	0.47	0.58	23%
	x tiempo de sesión y dificultad	cuadrática	0.16	0.18	13%

Comparativo

Variables	Aumento % (promedio)
Benjamín	50.58%
Carlos Abel	40.02%
Carlos Enrique	40.55%
Concepción	49.31%
Denisse	25.30%

General	41.15%
Modelo más utilizado	cuadrático

Conclusiones



- Capturan relaciones complejas: Los modelos no lineales pueden identificar interacciones y patrones que los modelos lineales no detectan, mejorando la precisión en contextos con estructuras de datos más complejas.
- Mayor flexibilidad, menor interpretación: Aunque suelen tener mejor desempeño, los modelos no lineales pueden ser más difíciles de interpretar, lo que representa una diferencia entre precisión y transparencia.



Gracias