

ANÁLISIS ESTADÍSTICO DEL AGOTAMIENTO EN TRABAJO REMOTO MEDIANTE MODELOS DE REGRESIÓN

LABORATORIO DE APRENDIZAJE ESTADÍSTICO

JULIETA MADRIGAL FLORES

GIBRÁN LEONARDO CHÁVEZ GONZÁLEZ

DIANA FERNANDA BARBOSA DUEÑAS

OBJETIVOS

General

Analizar los factores asociados al burnout en trabajo remoto mediante modelos de regresión lineal.

Específico

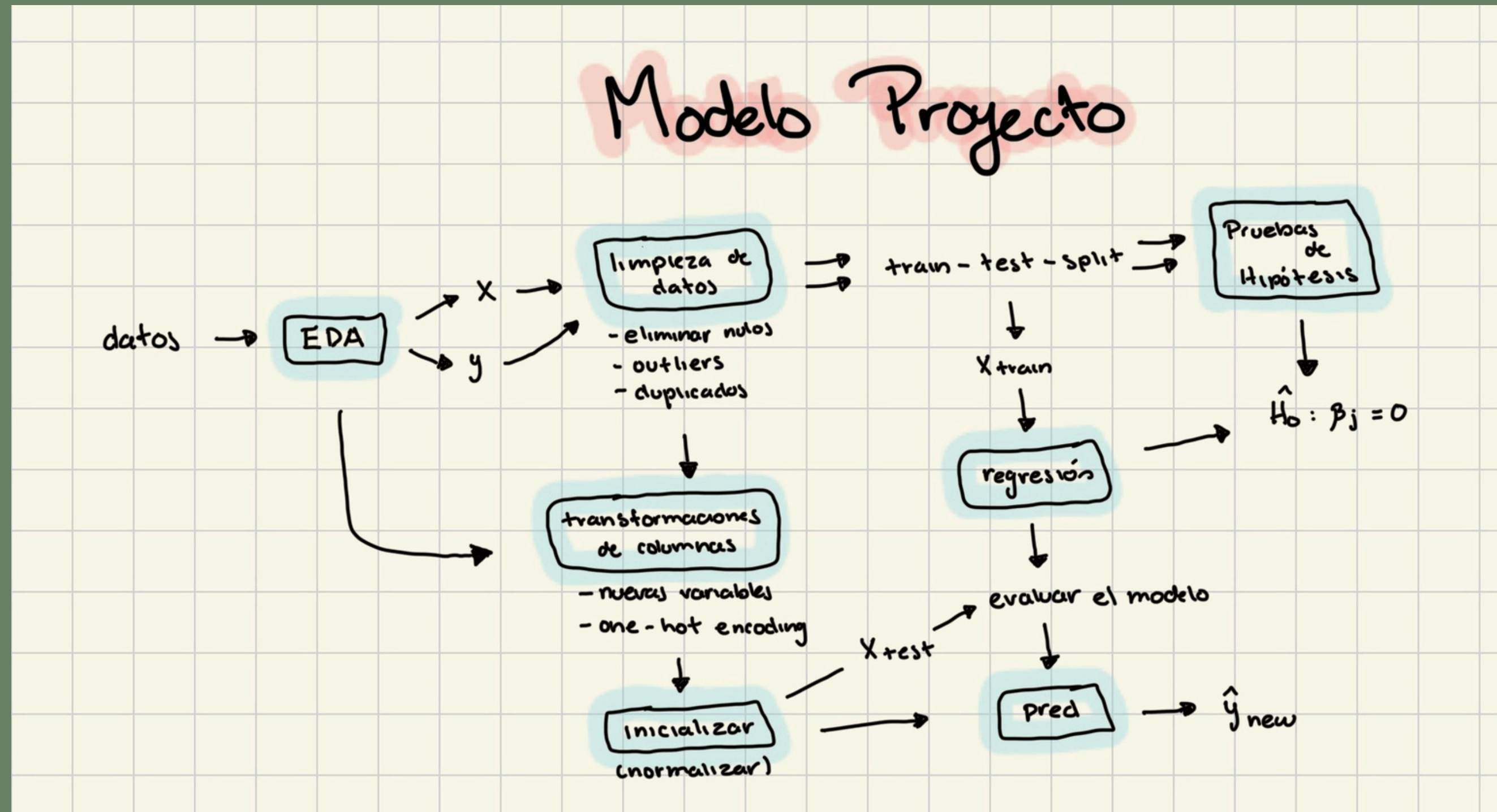
- Construir tres modelos de regresión distintos para explorar la relación entre las variables predictoras y el burnout score
- Evaluar la significancia estadística de cada variable en los modelos
- Comparar el desempeño de los modelos con y sin regularización

DATASET & EDA

	day_type	work_hours	screen_time_hours	meetings_count	breaks_taken	after_hours_work	sleep_hours	task_completion_rate
0	Weekday	9.59	11.86	4	2	0	7.55	91.2
1	Weekend	7.38	10.33	4	1	0	6.69	82.0
2	Weekend	6.31	8.92	1	2	0	8.87	80.6
3	Weekday	8.34	10.70	4	1	1	8.13	70.0
4	Weekend	6.97	9.83	1	2	0	5.85	67.1

	day_type	work_hours	screen_time_hours	meetings_count	breaks_taken	after_hours_work	sleep_hours	task_completion_rate
1		9.59	11.86	4	2	0	7.55	91.2
0		7.38	10.33	4	1	0	6.69	82.0
0		6.31	8.92	1	2	0	8.87	80.6
1		8.34	10.70	4	1	1	8.13	70.0
0		6.97	9.83	1	2	0	5.85	67.1

PIPELINE



MODELO 1

$$y = \beta_0 + \beta_1(\text{day_type}) + \beta_2(\text{work_hours}) + \beta_3(\text{screen_time_hours}) + \beta_4(\text{meetings_count}) + \beta_5(\text{breaks_taken}) + \beta_6(\text{after_hours_work}) + \beta_7(\text{sleep_hours}) + \beta_8(\text{task_completion_rate})$$

Variable	p-value
const	0.0000
day_type	0.3796
work_hours	0.1137
screen_time_hours	0.1042
meetings_count	0.2956
breaks_taken	0.6272
after_hours_work	0.2111
sleep_hours	0.4801
task_completion_rate	0.0000

	R ²
Sin penalización	0.919685
Lasso	0.748598
Ridge	0.919623
Elastic Net	0.226949

MODELO 2

$$y = \beta_0 + \beta_1(\text{after_hours_work}) + \beta_2(\text{task_completion_rate}) + \beta_3(\text{after_hours_work} \times \text{task_completion_rate})$$

Variable	p-value
const	0.0000
x1	0.4939
x2	0.0000
x3	0.6537

	R ²
Sin penalización	0.919382
Lasso	0.746377
Ridge	0.919297
Elastic Net	0.224318

MODELO 3

$$y = \beta_0 + \beta_1(\text{task_completion_rate}) + \beta_2(\text{task_completion_rate})^2$$

Variable	p-value
const	0.0
task_completion_rate	0.0
task_completion_rate^2	0.0

Variable	p-value
const	0.0
x1	0.0
x2	0.0

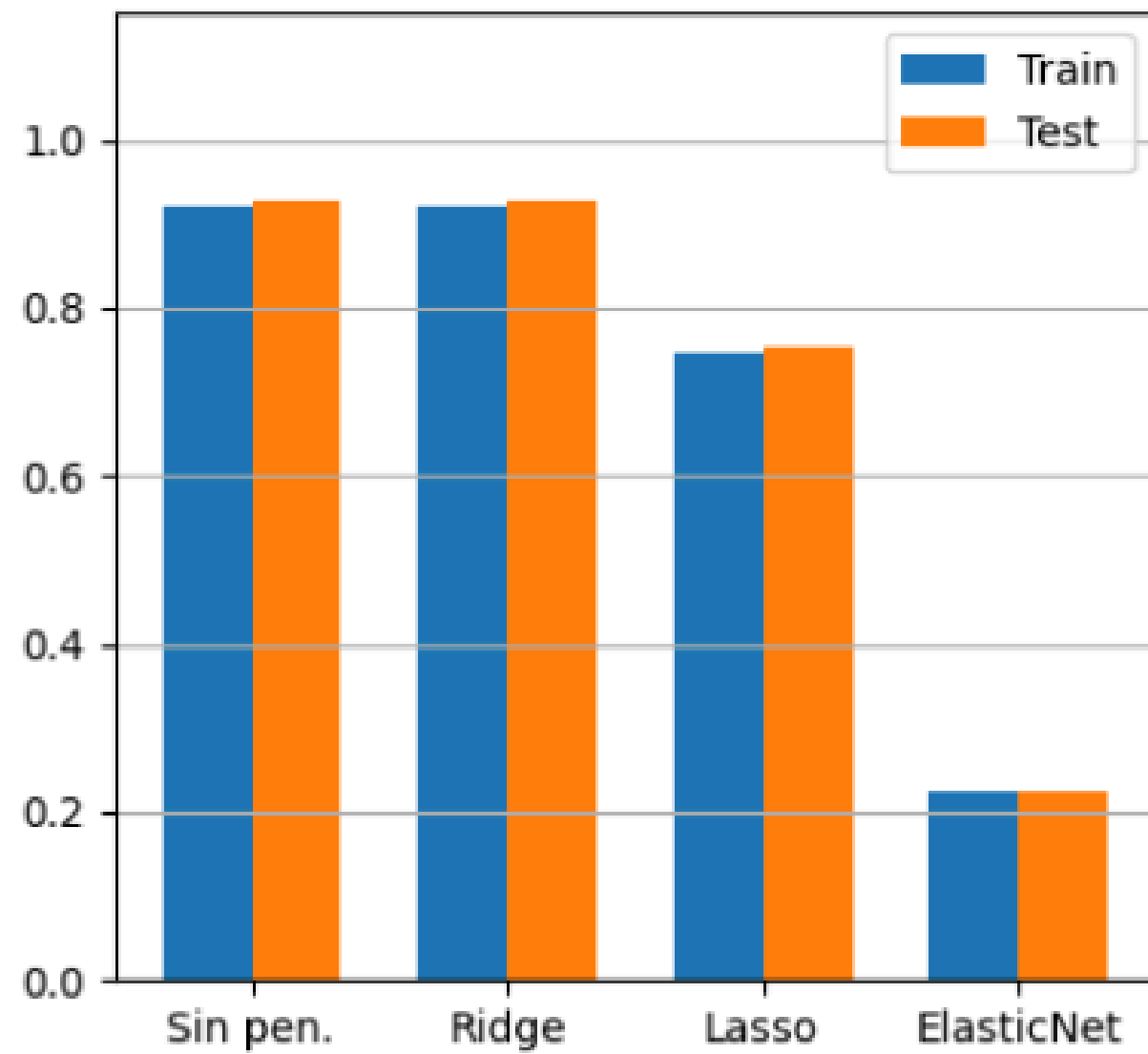
	R ²
Sin penalización	0.930028
Lasso	0.917033
Ridge	0.929769
Elastic Net	0.861010

TABLA R2

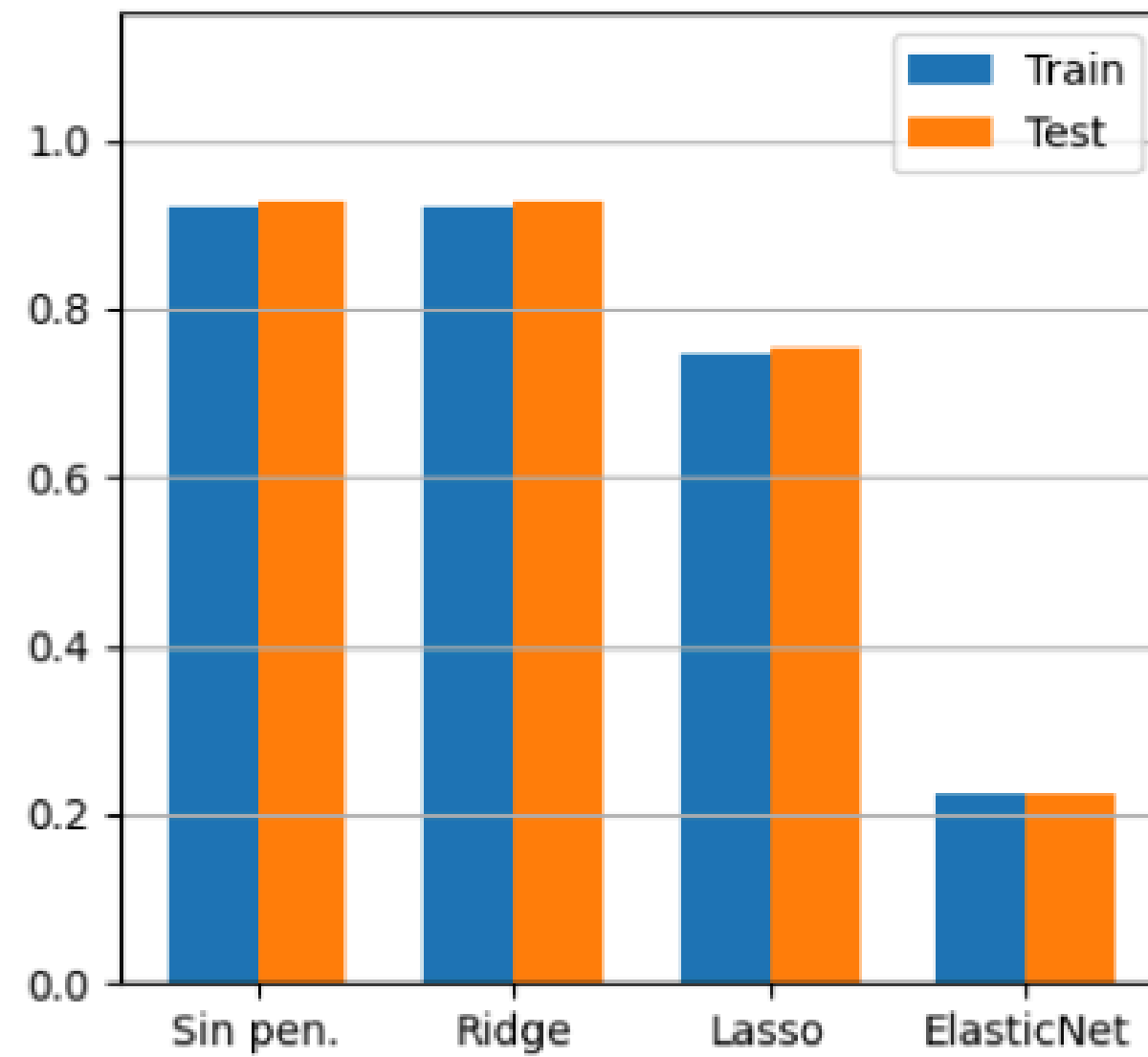
	M1 Train	M1 Test	M2 Train	M2 Test	M3 Train	M3 Test
Sin penalización	0.9197	0.9289	0.9194	0.9290	0.9300	0.9374
Ridge	0.9196	0.9289	0.9193	0.9290	0.9298	0.9375
Lasso	0.7464	0.7525	0.7464	0.7525	0.9170	0.9265
Elastic Net	0.2243	0.2245	0.2243	0.2245	0.8610	0.8696

VISUALIZACIÓN

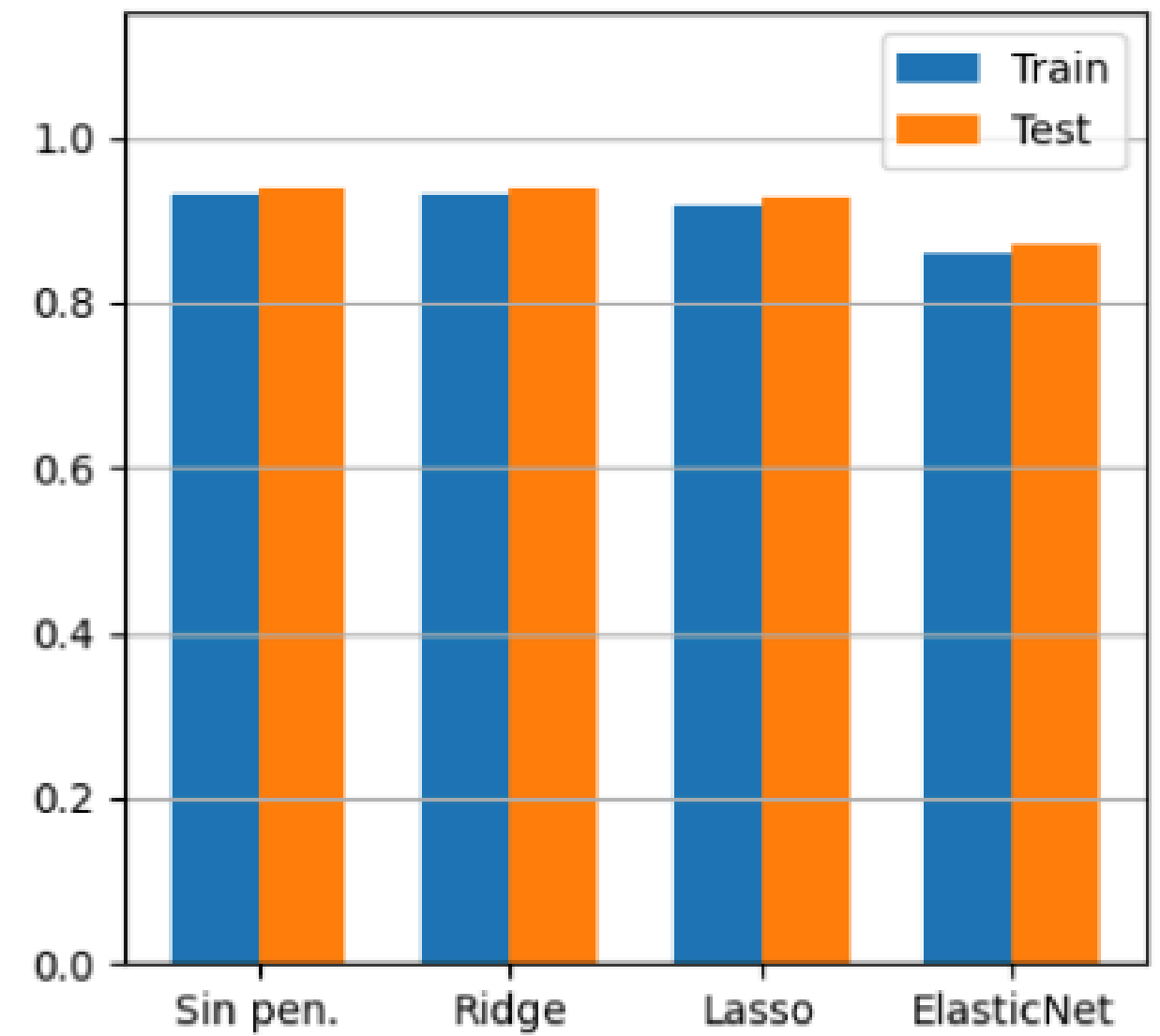
Modelo 1



Modelo 2



Modelo 3



CONCLUSIONES

- task_completion_rate es el único predictor estadísticamente significativo del burnout score
- La relación no lineal = tanto niveles bajos como altos de productividad se asocian con mayor agotamiento
- El burnout no depende de cuánto se trabaja, sino de qué se complete lo que se tiene pendiente
- La regularización no mejoró ningún modelo , el problema es la dominancia de una sola variable, no el sobreajuste
- Futuros estudios = incorporar variables como calidad del sueño, tipo de tareas o autonomía laboral para enriquecer el modelo

¡MUCHAS GRACIAS!