



ESCUELA DE INGENIERÍA
FACULTAD DE INGENIERÍA

EDUCACIÓN
PROFESIONAL

Fundamentos Machine & Deep Learning

Diplomado Big Data
y Analítica de Datos 2022-2023

Profesor:

Rodrigo Sandoval U.



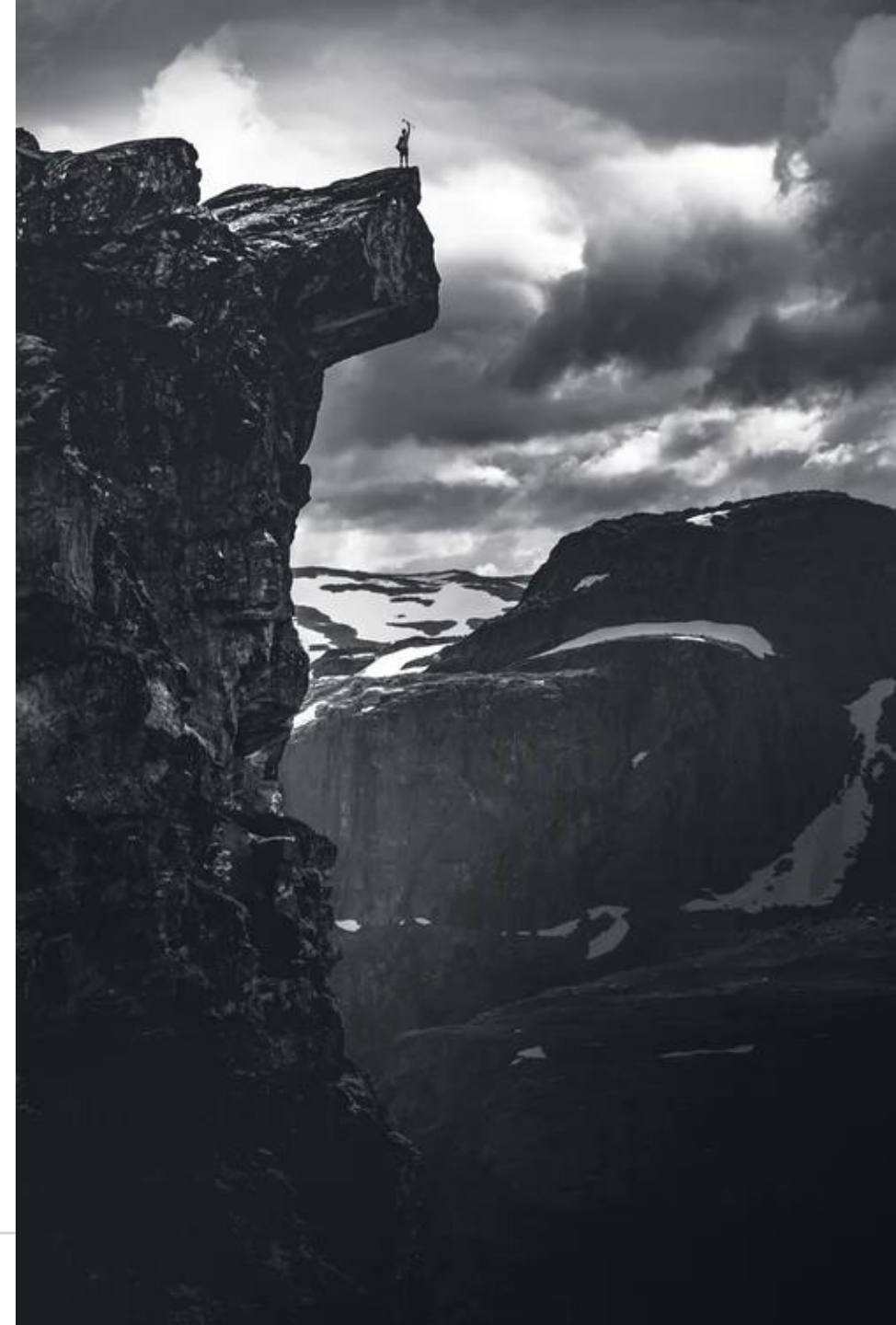


Mamá:

"¿Si tus compañeros saltan al precipicio, tú también saltas?"

Hijo Machine Learning:

"Por supuesto"



En Machine Learning Supervisado...

... los **datos** son los que hablan y cuentan la verdadera historia del fenómeno o situación, liberando a los humanos de tener que proponer las reglas lógicas de decisión.



Modelos de Machine Learning

PREDICCIÓN

Aprendizaje Supervisado

Clasificación

Árboles Decisión

Random Forest

K-Nearest Neighbor

SVM

Naïve Bayes

Redes Neuronales

Regresión

Lineal

No-Lineal

Logística

DESCRIPCIÓN

Aprendizaje No-supervisado

Asociación

Reglas de Asoc

Apriori

Segmentación

K-Means

Jerárquicos

DBSCAN

HDBSCAN

ANÁLISIS

Semi-supervisado

Descripción

Word Embedding

Autoencoders

Generative Adversarial Networks (GAN)

Transformers

# Partido	Tipo Cancha	Tiempo	Resultado
1	Arcilla	Soleado	GANÓ
2	Arcilla	Soleado	GANÓ
3	Césped	Nublado	PERDIÓ
4	Sintética	Nublado	GANÓ
5	Arcilla	Nublado	GANÓ
6	Sintética	Soleado	PERDIÓ
7	Arcilla	Nublado	GANÓ
8	Césped	Soleado	PERDIÓ
9	Arcilla	Soleado	PERDIÓ
10	Sintética	Nublado	GANÓ
11	Sintética	Soleado	GANÓ
12	Sintética	Nublado	GANÓ



Siguiente Partido → ¿Gana o Pierde?

Partido 13: Césped + Soleado → ¿Gana o Pierde?

Partido 14: Arcilla + Soleado → ¿Gana o Pierde?

Partido 15: Sintético + Nublado → ¿Gana o Pierde?

Calcular la Probabilidad de Ganar

	Arcilla	Césped	Sintético
Soleado	2 1	0 1	1 0
Nublado	2 0	0 1	3 1

Partido 13: Césped + Soleado → ¿Gana o Pierde?

Partido 14: Arcilla + Soleado → ¿Gana o Pierde?

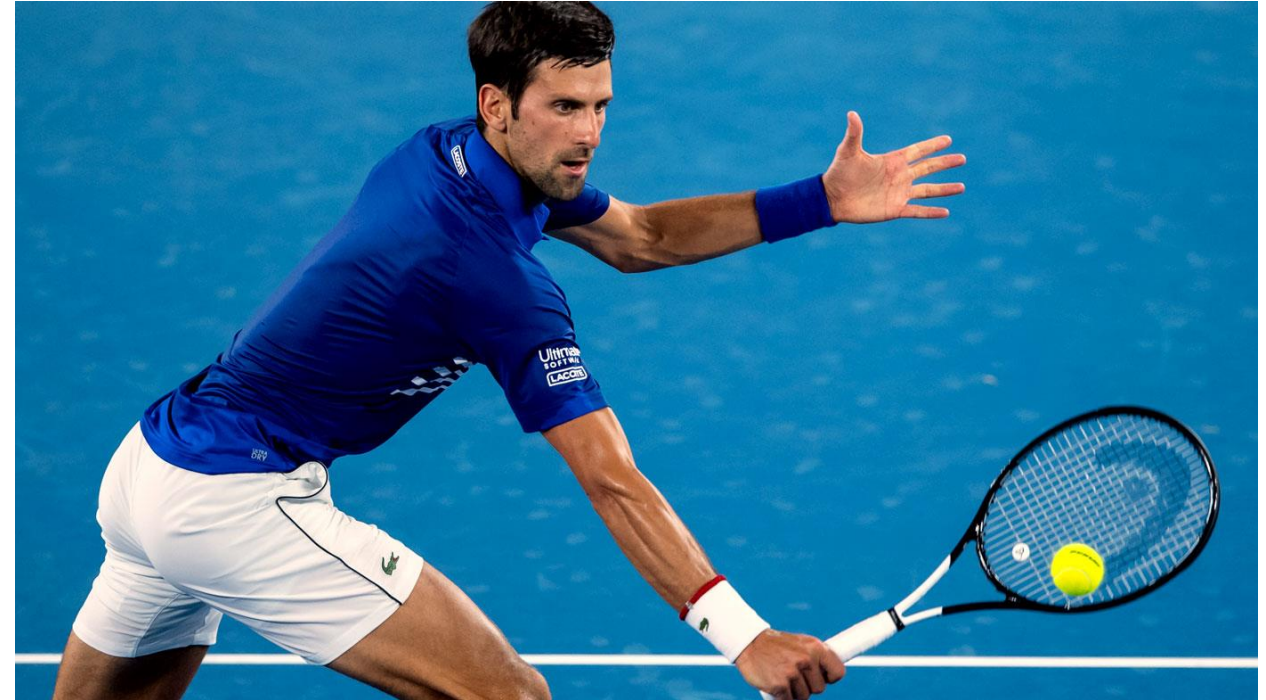
Partido 15: Sintético + Nublado → ¿Gana o Pierde?

Análisis del ejemplo ...

Preguntas Esenciales

¿Soleado/Nublado y Tipo de Cancha son los únicos parámetros relevantes de analizar?

¿Qué otros podrían ser relevantes?



Entonces, si en vez de 12 partidos y 2 atributos tuviésemos mucho más ...

... es muy probable que podríamos hacer mejores predicciones: más precisas, más reales.

Podríamos resolver necesidades de negocio.

#	Tipo Cancha	Tiempo	R.Op	Desc	Alim	Resultado
1	Arcilla	Soleado	10	2	Bien	GANÓ
2	Arcilla	Soleado	14	3	Regular	GANÓ

#	Tipo Cancha	Tiempo	R.Op	Desc	Alim	Resultado
3	Césped	Nublado	10	2	Bien	GANÓ
4	Sintética	Nublado	14	3	Regular	GANÓ
5	Arcilla	Soleado	22	7	Bien	PERDIÓ
6	Sintética	Soleado	-13	12	Poco	GANÓ
7	Arcilla	Nublado	-14	1	Bien	GANÓ
8	Césped	Soleado	8	1	Regular	PERDIÓ
9	Arcilla	Soleado	6	2	Bien	GANÓ
10	Sintética	Nublado	11	5	Poco	PERDIÓ
11	Sintética	Soleado	15	4	Bien	PERDIÓ
12	Arcilla	Soleado	-5	17	Regular	GANÓ
13	Sintética	Nublado	-6	18	Bien	GANÓ
14	Arcilla	Soleado	10	2	Poco	GANÓ
15	Césped	Nublado	14	3	Bien	GANÓ
16	Sintética	Nublado	22	7	Regular	GANÓ
17	Arcilla	Soleado	-13	12	Bien	PERDIÓ
18	Sintética	Nublado	-14	1	Poco	GANÓ
19	Arcilla	Nublado	8	1	Bien	GANÓ
20	Césped	Nublado	6	2	Regular	PERDIÓ
21	Arcilla	Soleado	-5	17	Bien	PERDIÓ
22	Sintética	Nublado	-6	18	Regular	GANÓ
23	Arcilla	Soleado	10	2	Bien	GANÓ
24	Césped	Nublado	14	3	Poco	GANÓ
25	Arcilla	Soleado	22	7	Bien	GANÓ
26	Arcilla	Soleado	-13	12	Regular	GANÓ
27	Césped	Nublado	-14	1	Bien	PERDIÓ
28	Arcilla	Nublado	8	1	Poco	GANÓ
29	Arcilla	Nublado	6	2	Bien	GANÓ
30	Sintética	Soleado	11	5	Regular	PERDIÓ
31	Arcilla	Nublado	15	4	Bien	GANÓ
32	Césped	Soleado	-5	17	Poco	PERDIÓ
33	Arcilla	Soleado	-6	18	Bien	PERDIÓ
34	Sintética	Nublado	10	2	Regular	GANÓ
35	Arcilla	Soleado	14	3	Bien	GANÓ
36	Sintética	Nublado	22	7	Poco	GANÓ
37	Arcilla	Soleado	-13	12	Bien	GANÓ
38	Arcilla	Soleado	-14	1	Regular	GANÓ
39	Césped	Nublado	8	1	Bien	PERDIÓ
40	Sintética	Nublado	6	2	Poco	GANÓ

#	Tipo Cancha	Tiempo	R.Op	Desc	Alim	Resultado
1	Arcilla	Soleado	10	2	Bien	GANÓ
2	Arcilla	Soleado	14	3	Regular	GANÓ
3	Césped	Nublado	22	7	Bien	PERDIÓ
4	Sintética	Nublado	-13	12	Poco	GANÓ
5	Arcilla	Soleado	-14	1	Bien	GANÓ
6	Césped	Soleado	8	1	Regular	PERDIÓ
7	Arcilla	Soleado	6	2	Bien	GANÓ
8	Sintética	Nublado	11	5	Poco	PERDIÓ
9	Arcilla	Soleado	15	4	Bien	PERDIÓ
10	Sintética	Nublado	-5	17	Regular	GANÓ
11	Arcilla	Soleado	-6	18	Bien	GANÓ
12	Césped	Nublado	10	2	Poco	GANÓ
13	Arcilla	Soleado	14	3	Bien	GANÓ
14	Sintética	Nublado	22	7	Regular	GANÓ
15	Arcilla	Soleado	-13	12	Bien	PERDIÓ
16	Sintética	Nublado	-14	1	Poco	GANÓ
17	Arcilla	Nublado	8	1	Bien	GANÓ
18	Césped	Nublado	6	2	Regular	PERDIÓ
19	Arcilla	Soleado	-5	17	Bien	PERDIÓ
20	Sintética	Nublado	-6	18	Regular	GANÓ
21	Arcilla	Soleado	10	2	Bien	GANÓ
22	Césped	Nublado	14	3	Poco	GANÓ
23	Arcilla	Soleado	22	7	Bien	GANÓ
24	Sintética	Nublado	-13	12	Bien	PERDIÓ
25	Arcilla	Soleado	-14	1	Poco	GANÓ
26	Césped	Nublado	8	1	Bien	GANÓ
27	Arcilla	Soleado	6	2	Regular	PERDIÓ
28	Sintética	Nublado	-5	17	Bien	PERDIÓ
29	Arcilla	Soleado	-6	18	Regular	GANÓ
30	Arcilla	Soleado	10	2	Bien	GANÓ
31	Césped	Nublado	14	3	Poco	GANÓ
32	Arcilla	Soleado	22	7	Bien	GANÓ
33	Sintética	Nublado	-13	12	Regular	GANÓ
34	Arcilla	Soleado	-14	1	Bien	PERDIÓ
35	Césped	Nublado	8	1	Poco	GANÓ
36	Arcilla	Soleado	6	2	Bien	GANÓ
37	Sintética	Nublado	11	5	Regular	PERDIÓ
38	Arcilla	Soleado	15	4	Bien	GANÓ
39	Césped	Soleado	-5	17	Poco	PERDIÓ
40	Arcilla	Soleado	-6	18	Bien	PERDIÓ
41	Sintética	Nublado	10	2	Regular	GANÓ
42	Arcilla	Soleado	14	3	Bien	GANÓ
43	Sintética	Nublado	22	7	Poco	GANÓ
44	Arcilla	Soleado	-13	12	Bien	GANÓ
45	Sintética	Nublado	-14	1	Regular	GANÓ
46	Césped	Nublado	8	1	Bien	PERDIÓ
47	Arcilla	Soleado	6	2	Poco	GANÓ
48	Sintética	Nublado	11	5	Regular	PERDIÓ
49	Arcilla	Soleado	-5	17	Poco	PERDIÓ
50	Césped	Nublado	-6	18	Bien	PERDIÓ
51	Arcilla	Soleado	10	2	Regular	GANÓ
52	Sintética	Nublado	14	3	Bien	GANÓ
53	Césped	Soleado	22	7	Poco	GANÓ
54	Arcilla	Soleado	-13	12	Bien	GANÓ
55	Sintética	Nublado	-14	1	Regular	GANÓ
56	Césped	Nublado	8	1	Bien	PERDIÓ
57	Arcilla	Soleado	6	2	Poco	GANÓ
58	Sintética	Nublado	11	5	Regular	PERDIÓ
59	Arcilla	Soleado	-5	17	Poco	PERDIÓ
60	Césped	Nublado	-6	18	Bien	PERDIÓ
61	Arcilla	Soleado	10	2	Regular	GANÓ
62	Sintética	Nublado	14	3	Bien	GANÓ
63	Césped	Soleado	22	7	Poco	GANÓ
64	Arcilla	Soleado	-13	12	Bien	GANÓ
65	Sintética	Nublado	-14	1	Regular	GANÓ
66	Césped	Nublado	8	1	Bien	PERDIÓ
67	Arcilla	Soleado	6	2	Poco	GANÓ
68	Sintética	Nublado	11	5	Regular	PERDIÓ
69	Arcilla	Soleado	-5	17	Poco	PERDIÓ
70	Césped	Nublado	-6	18	Bien	PERDIÓ
71	Arcilla	Soleado	10	2	Regular	GANÓ
72	Sintética	Nublado	14	3	Bien	GANÓ
73	Césped	Soleado	22	7	Poco	GANÓ
74	Arcilla	Soleado	-13	12	Bien	GANÓ
75	Sintética	Nublado	-14	1	Regular	GANÓ
76	Césped	Nublado	8	1	Bien	PERDIÓ
77	Arcilla	Soleado	6	2	Poco	GANÓ
78	Sintética	Nublado	11	5	Regular	PERDIÓ
79	Arcilla	Soleado	-5	17	Poco	PERDIÓ
80	Césped	Nublado	-6	18	Bien	PERDIÓ
81	Arcilla	Soleado	10	2	Regular	GANÓ
82	Sintética	Nublado	14	3	Bien	GANÓ
83	Césped	Soleado	22	7	Poco	GANÓ
84	Arcilla	Soleado	-13	12	Bien	GANÓ
85	Sintética	Nublado	-14	1	Regular	GANÓ
86	Césped	Nublado	8	1	Bien	PERDIÓ
87	Arcilla	Soleado	6	2	Poco	GANÓ
88	Sintética	Nublado	11	5	Regular	PERDIÓ
89	Arcilla	Soleado	-5	17	Poco	PERDIÓ
90	Césped	Nublado	-6	18	Bien	PERDIÓ
91	Arcilla	Soleado	10	2	Regular	GANÓ
92	Sintética	Nublado	14	3	Bien	GANÓ
93	Césped	Soleado	22	7	Poco	GANÓ
94	Arcilla	Soleado	-13	12	Bien	GANÓ
95	Sintética	Nublado	-14	1	Regular	GANÓ
96	Césped	Nublado	8	1	Bien	PERDIÓ
97	Arcilla	Soleado	6	2	Poco	GANÓ
98	Sintética	Nublado	11	5	Regular	PERDIÓ
99	Arcilla	Soleado	-5	17	Poco	PERDIÓ
100	Césped	Nublado	-6	18	Bien	PERDIÓ



Cambiamos el tipo de pregunta

# Partido	Tipo Cancha	Tiempo	Duración
1	Arcilla	Soleado	2 horas
2	Arcilla	Soleado	2,5 horas
3	Césped	Nublado	1,5 horas
4	Sintética	Nublado	3 horas
5	Arcilla	Nublado	2 horas
6	Sintética	Soleado	3 horas
7	Arcilla	Nublado	2 horas
8	Césped	Soleado	1,5 horas
9	Arcilla	Soleado	2,5 horas
10	Sintética	Nublado	3,5 horas
11	Sintética	Soleado	2,5 horas
12	Sintética	Nublado	3 horas



Siguiente Partido → ¿Cuánto dura?

Partido 13: Césped + Soleado → ¿Cuánto dura?

Partido 14: Arcilla + Soleado → ¿Cuánto dura?

Partido 15: Sintético + Nublado → ¿Cuánto dura?



Cálculos estadísticos para poder responder

Promedio(Todos)

- 2,42 horas

Promedio (Césped)

- 1,5 horas

Promedio (Arcilla)

- 2,2 horas

Promedio (Sintético)

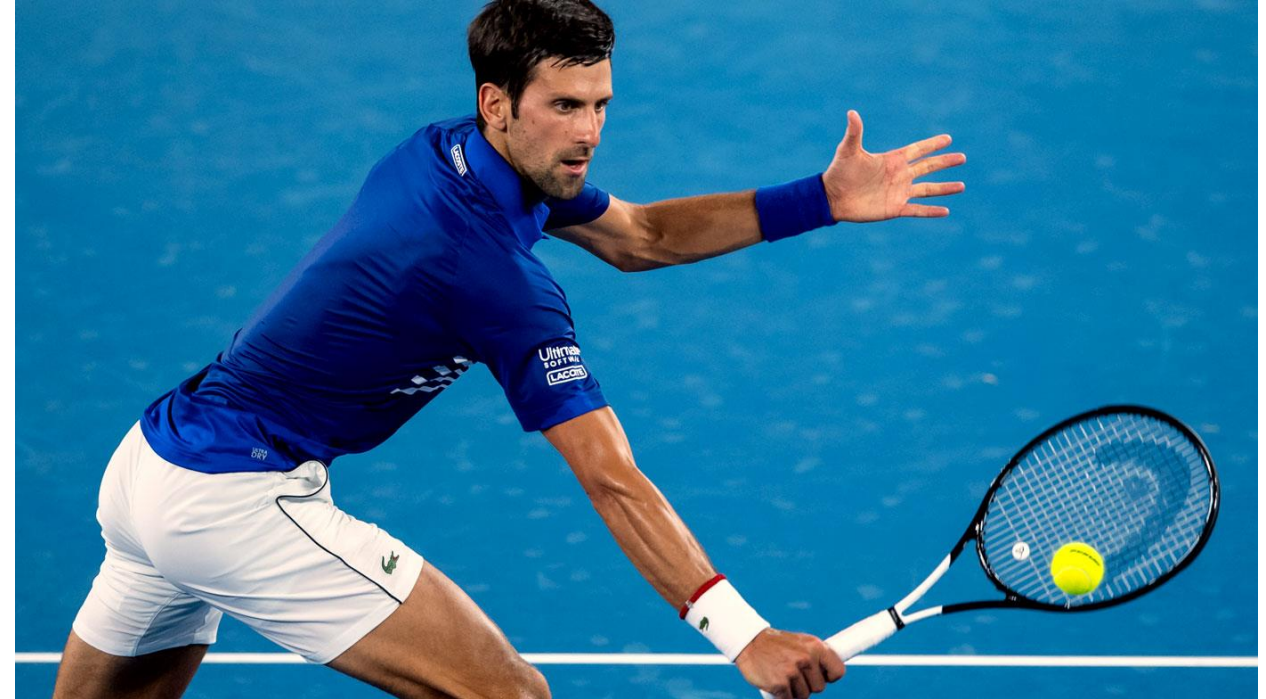
- 3,2 horas

Análisis del ejemplo ...

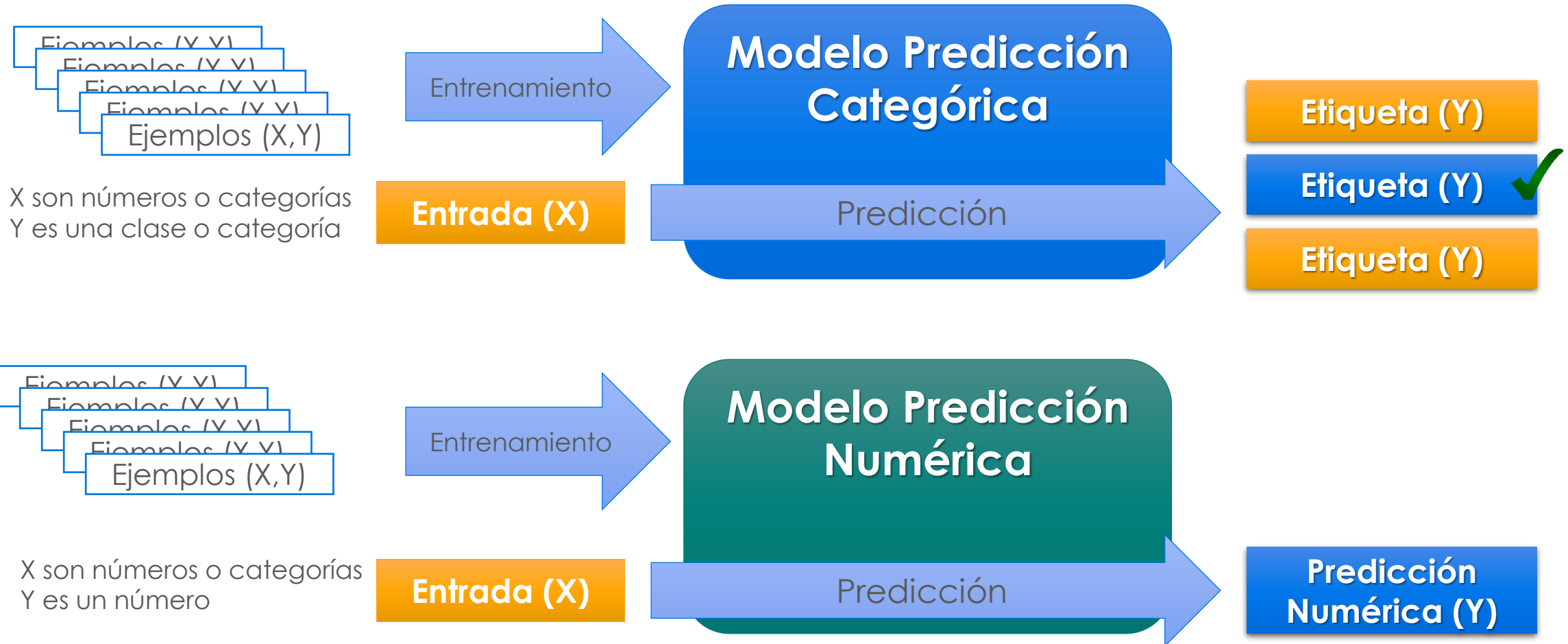
El cambio de pregunta implica y un cambio en el método de predicción.

La predicción se basa en los atributos que hay.

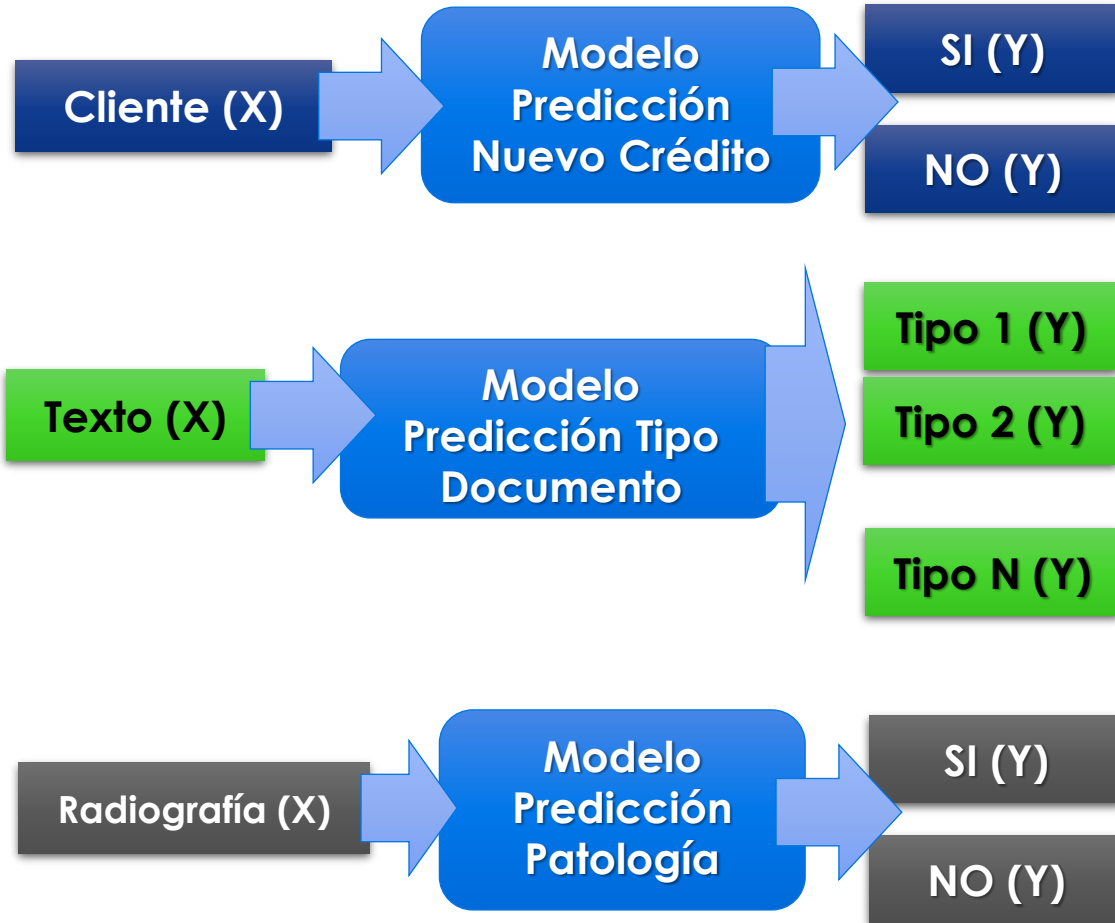
Resulta relevante buscar alguna correlación entre los datos conocidos y los resultados.



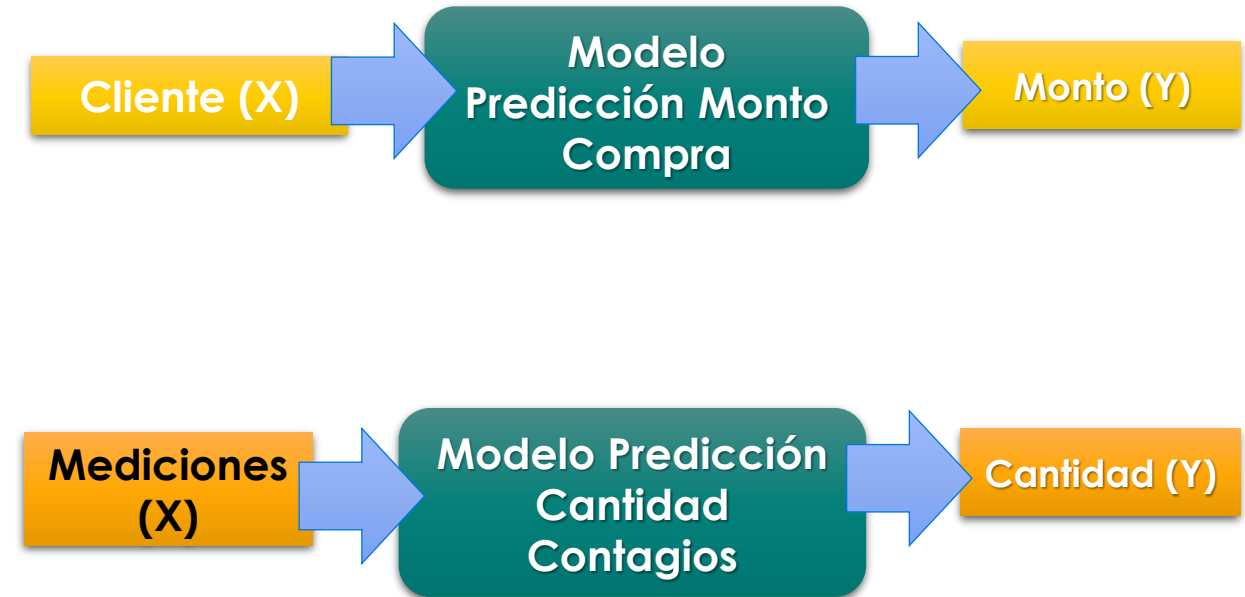
Resumen de técnicas de aprendizaje supervisado



Predicción Categórica o Clasificación Supervisada

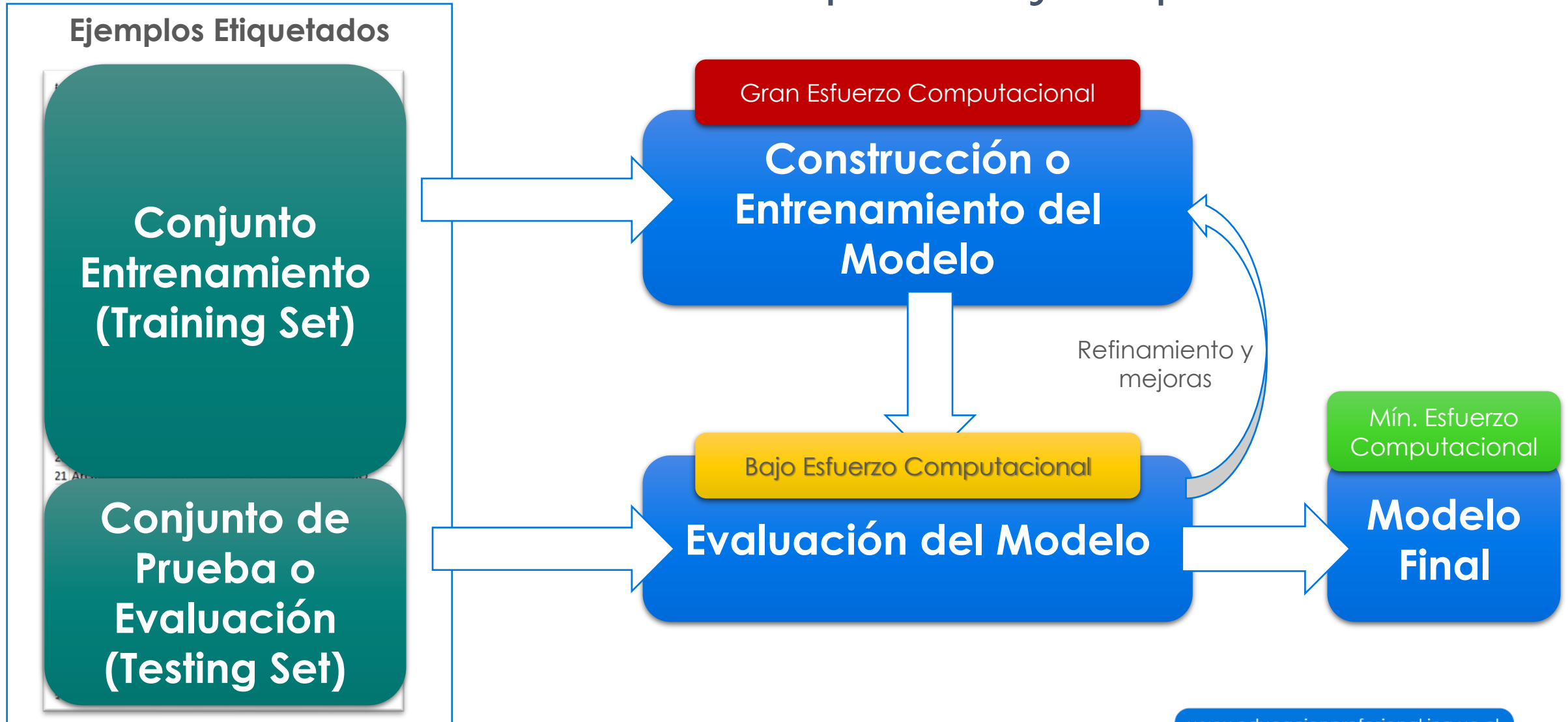


Predicción Numérica





Entrenamiento de Modelos de Aprendizaje Supervisado





Desafíos en Aprendizaje Supervisado

Disponibilidad
de Suficientes
Datos

Calidad de los
Datos

Cantidad de
Datos y
Capacidad de
Entrenamiento



Estrategias para mejorar desempeño de modelos

Limpieza
dataset

Balanceo
dataset

Métricas para
medir el
desempeño

Métricas de Desempeño de Modelos

Predicción Categórica

- Accuracy
- Sensitivity/Specificity
- Otros

Matriz de Confusión

Accuracy	$\frac{TP+TN}{P+N}$
Exactitud	

		Predicción			
		P	N		
Real	P	TP	FN	P	
	N	FP	TN		N
		P'	N'		

Predicción Numérica

- RMSE
- MAPE
- Otros

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2}$$

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{e_t}{y_t} \right|$$



Gracias



rsandova@ing.puc.cl
rodrigo@RSolver.com



@RSandovalSolver



/in/RodrigoSandoval

www.RodrigoSandoval.net
www.RSolver.com