



Introducción a Machine Learning

Carlos Adrián Alarcón



EXPERIENCIA

- Data Scientist en Belcorp
- CoFounder en DataLysis Perú
- Profesional especializado en proyectos de Business Analytics y Business Intelligence.
- MBA por CENTRUM PUCP, candidato a Master in Data Science por Universidad Adolfo Ibañez y Universidad Los Andes
- +4 años de experiencia en proyectos de Business Intelligence y Analytics en diversas empresas
- Instructor de SQL Server, Power BI, Python y R

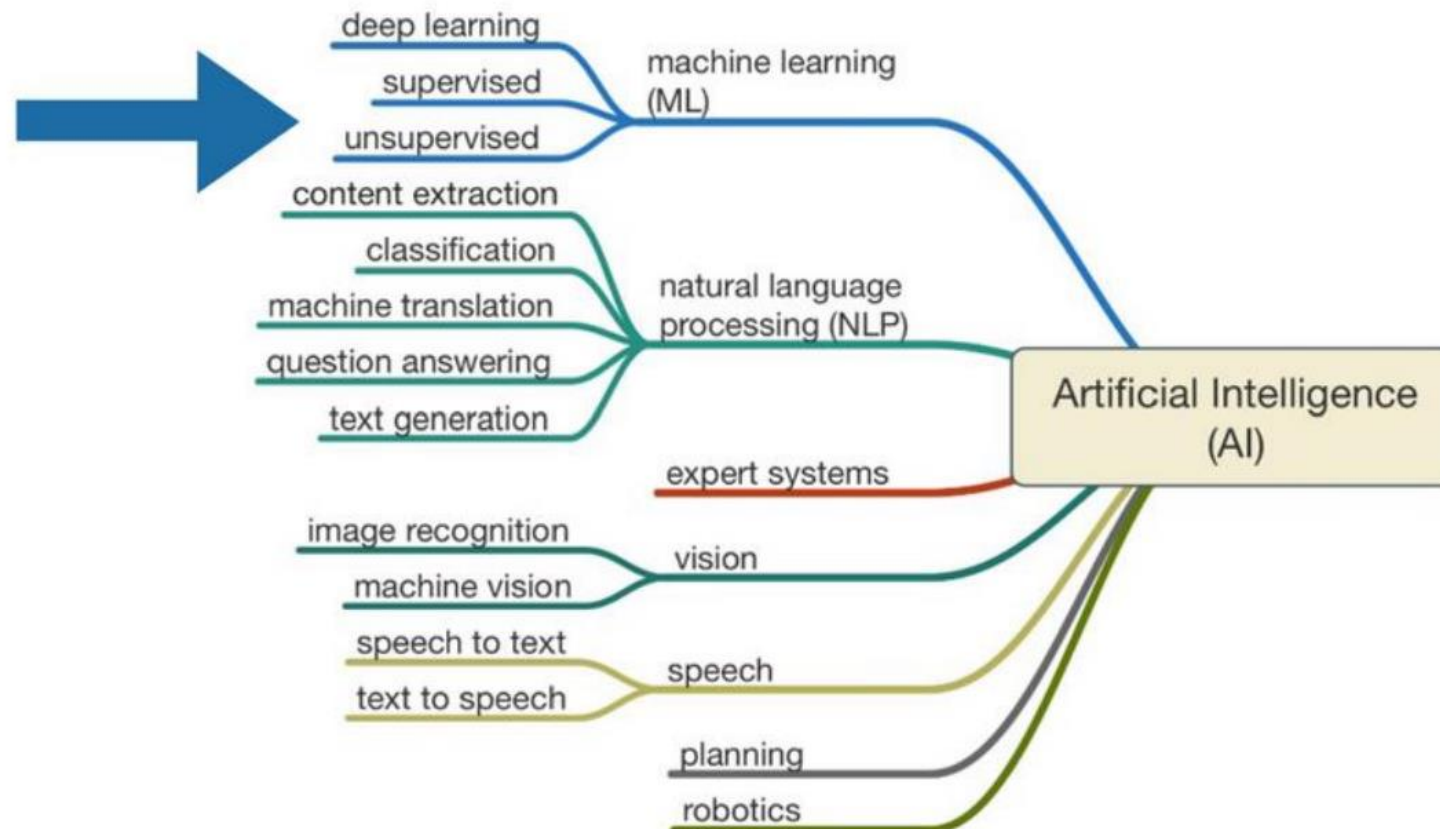
CERTIFICACIONES



Machine Learning



El aprendizaje automático comprende el estudio y diseño de programas que mejoran el desempeño P al realizar una tarea T , a partir de la experiencia E . (Tom Mitchell, 1997)



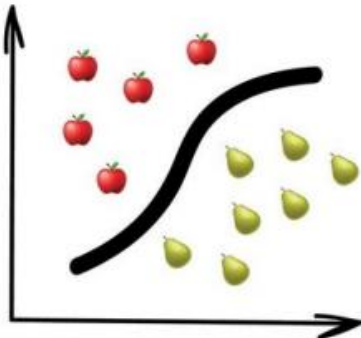
Machine Learning

Aprendizaje Supervisado

Datos etiquetados

Clasificación

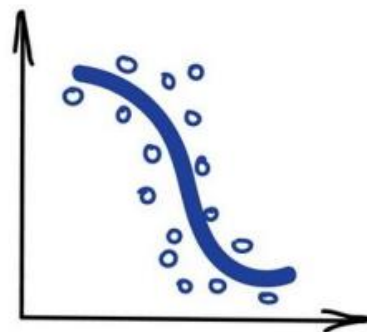
Predecir una categoría



Classification

Regresión

Predecir un número



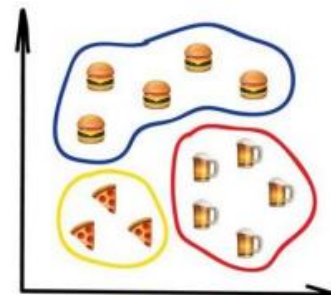
Regression

Aprendizaje No Supervisado

Datos no etiquetados

Agrupación

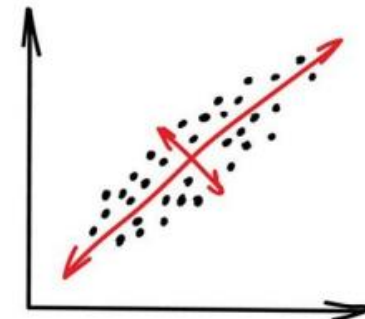
Dividir por similitud



Clustering

Reducción

Reducir dimensionalidad



Dimension Reduction

Dataprep for machine learning



Variables categóricas

¿Qué hacemos cuando tenemos variables categóricas?

One Hot Encoding vs Label Encoding

Label Encoding

Food Name	Categorical #	Calories
Apple	1	95
Chicken	2	231
Broccoli	3	50



One Hot Encoding

Apple	Chicken	Broccoli	Calories
1	0	0	95
0	1	0	231
0	0	1	50

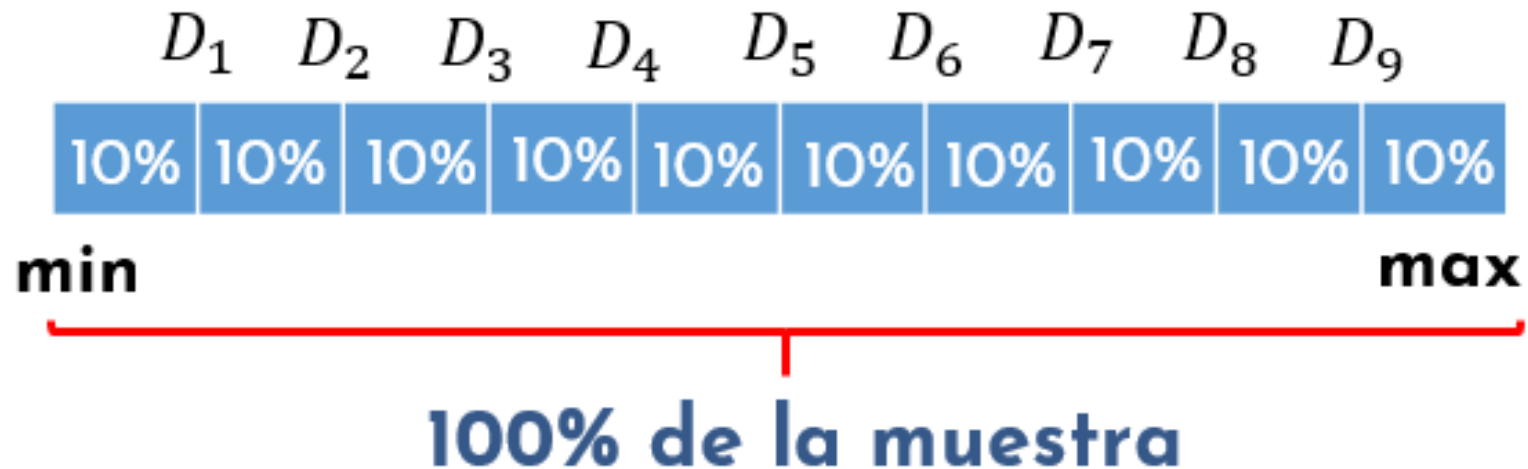
Dataprep for machine learning



Variables continuas

Y si pasa al revés ...

¿Cómo agrupamos variables continuas?

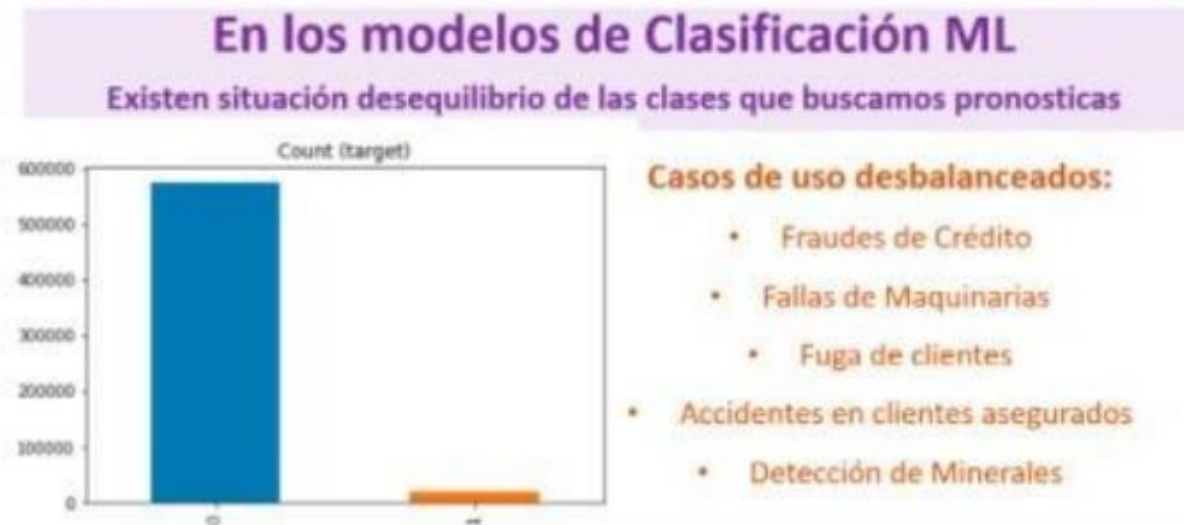


Dataprep for machine learning



Data desbalanceada

Si hay datos desbalanceados ... , ¿qué hacemos?



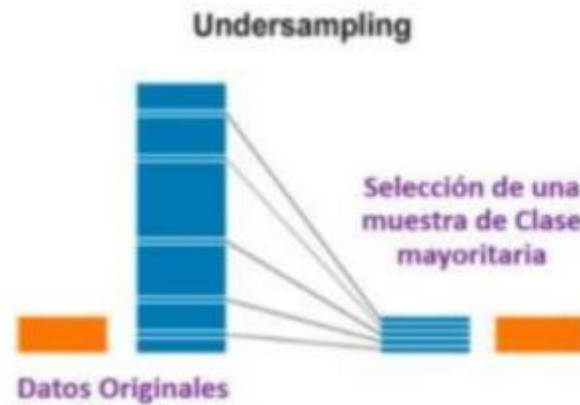
El desbalance de las clases puede afectar un poco los resultados del poder predictivos de algunos modelos

Dataprep for machine learning



Data desbalanceada

Si hay datos desbalanceados ... , ¿qué hacemos?



Dataprep for machine learning



Data desbalanceada

Si hay datos desbalanceados ... , ¿qué hacemos?

	Predicción Gato	Predicción Perro
Valor real Gato	Aciertos 990	0
Valor real Perro	Fallos 10	0

Accuracy	$\frac{990 + 0}{990 + 0 + 10 + 0}$	Precisión Clase 1 $\frac{990}{990 + 10}$	Recall Clase 1 $\frac{990}{990 + 0}$
		Precisión Clase 2 $\frac{0}{0 + 0}$	Recall Clase 2 $\frac{0}{0 + 10}$