

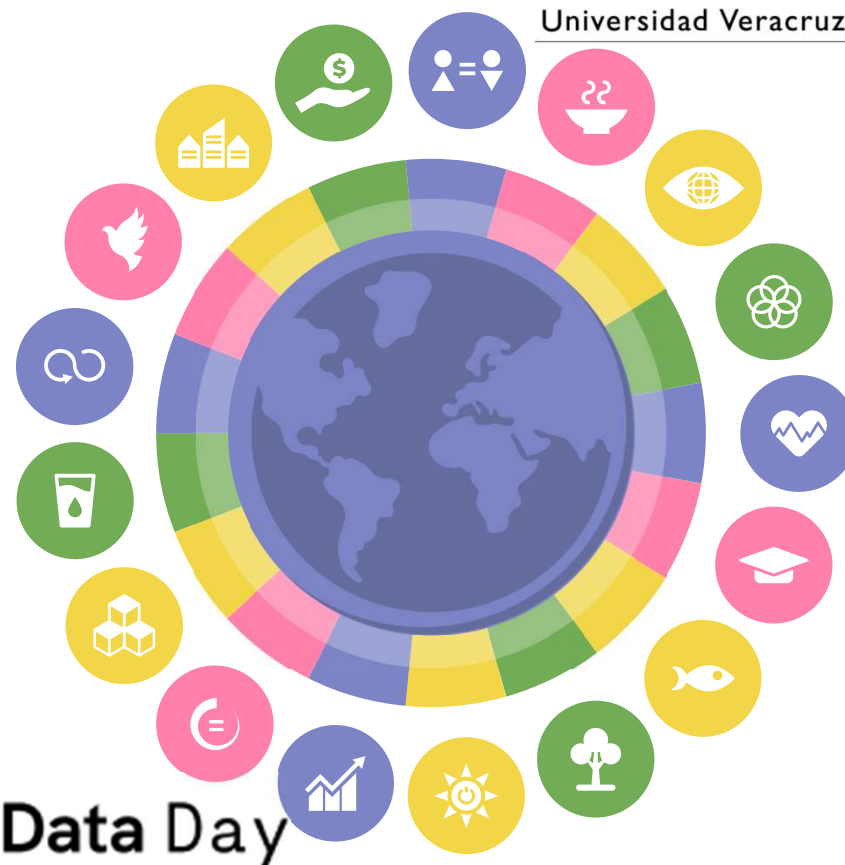
Datathon para la visualización de los ODS utilizando Datos Abiertos



Día de los Datos Abiertos, 2024



Universidad Veracruzana



Open Data Day





Programa



01

ODS de la Agenda 2030

MCD Lorena López Lozada

02

Introducción al manejo de datos

Dr. Ángel Juan Sánchez García

03

Visualización espacial de datos

Dr. Ángel Fernando Argüello Ortiz

04

Introducción al Aprendizaje Automático

MIA Juan Luis López Herrera

MIA. Juan Luis López Herrera



julopez@uv.mx



04

Aprendizaje Automático



Introducción



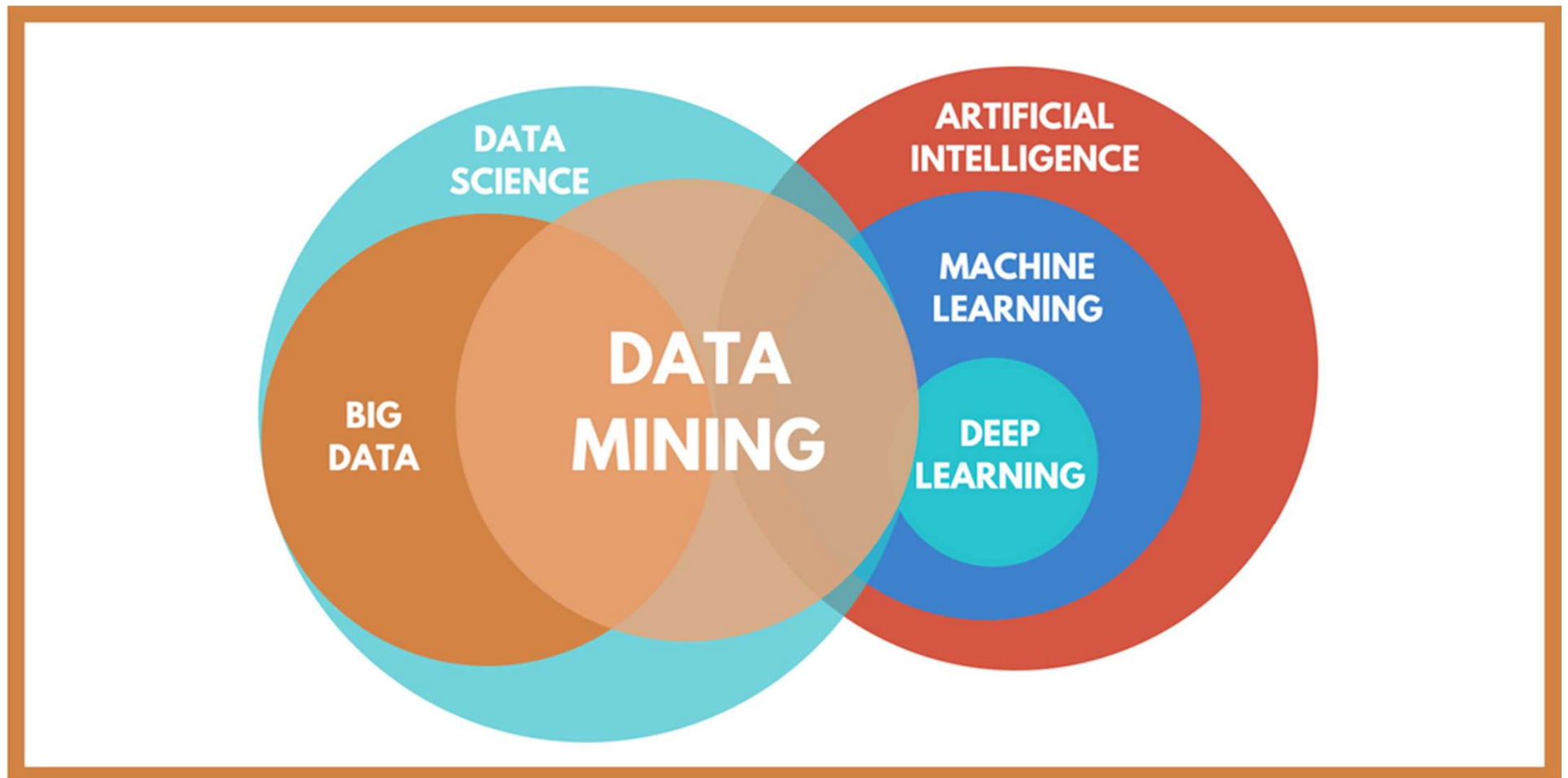
Origen and definición de los ODS

El **aprendizaje automático (AA)**; también llamado **automatizado, computacional o de máquinas** (del inglés machine learning, ML), es el subcampo de las **ciencias de la computación** y una **rama de la inteligencia artificial**, cuyo objetivo es desarrollar técnicas que permitan que las computadoras aprendan. Se dice que un agente aprende cuando su desempeño mejora con la experiencia y mediante el uso de datos; es decir, cuando la habilidad no estaba presente en su genotipo o rasgos de nacimiento

En el aprendizaje de máquinas un computador observa datos, construye un modelo basado en esos datos y utiliza ese modelo a la vez como una hipótesis acerca del mundo y una pieza de software que puede resolver problemas



Contexto del Aprendizaje automático



KDD (Knowledge Discovery in Databases)

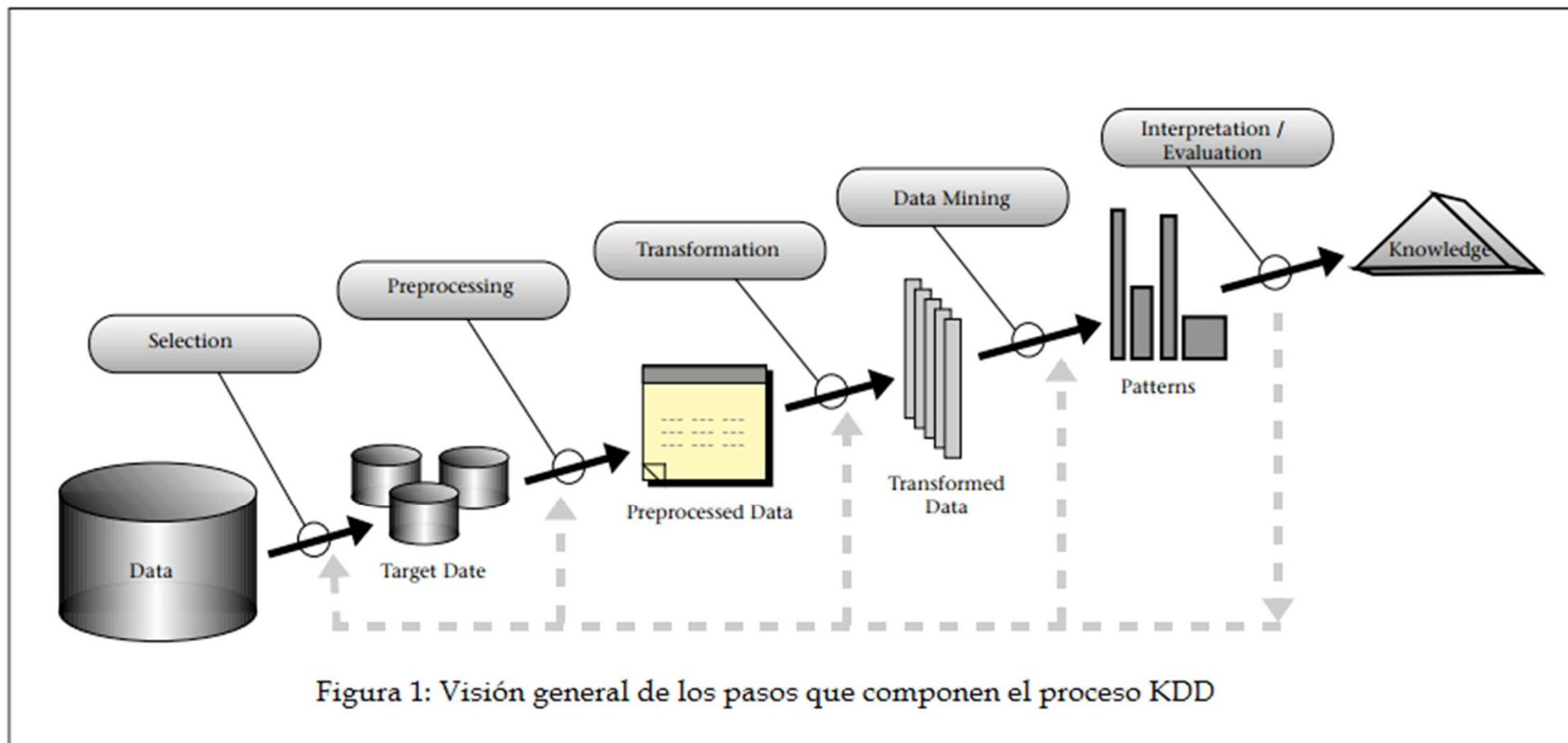
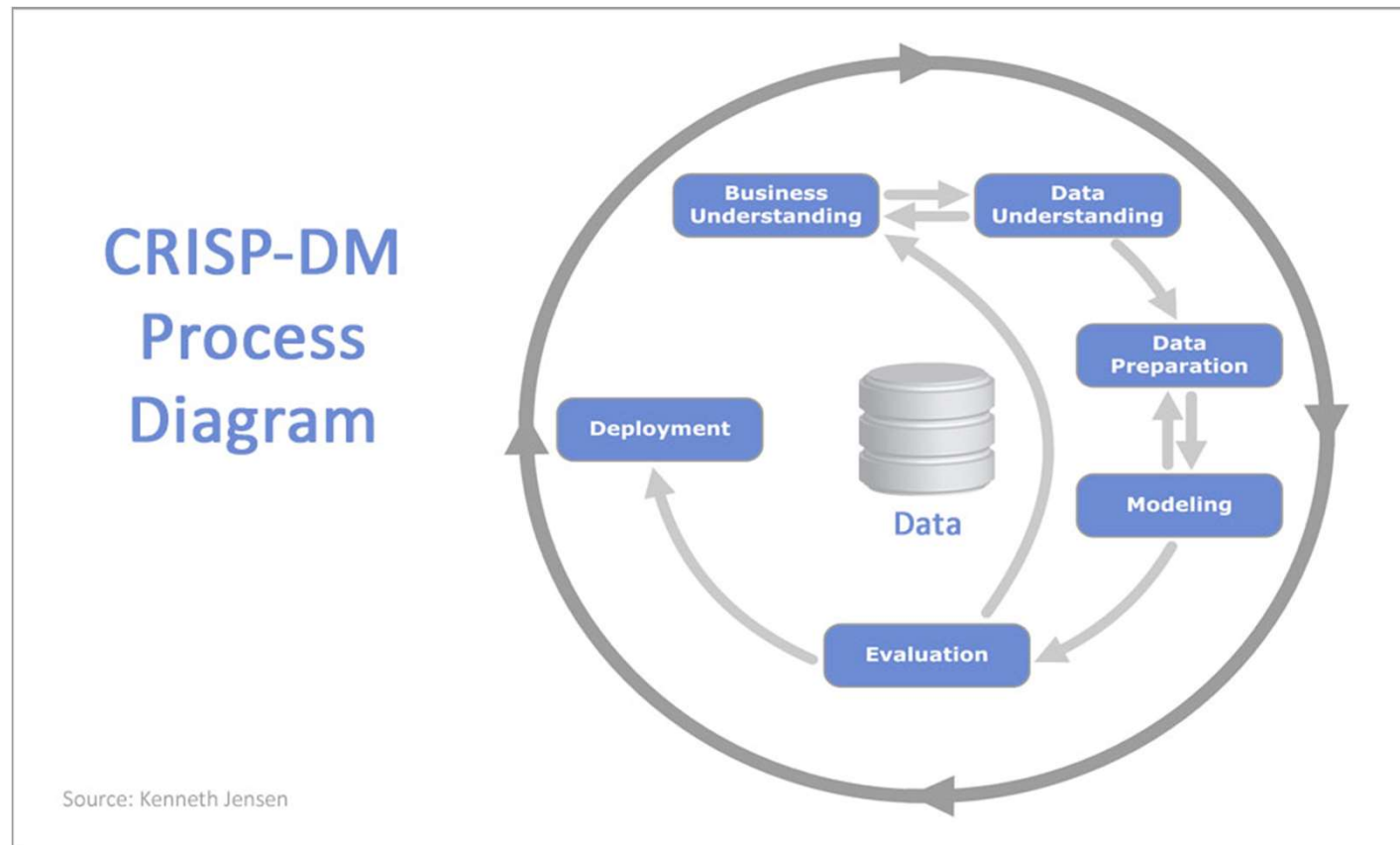


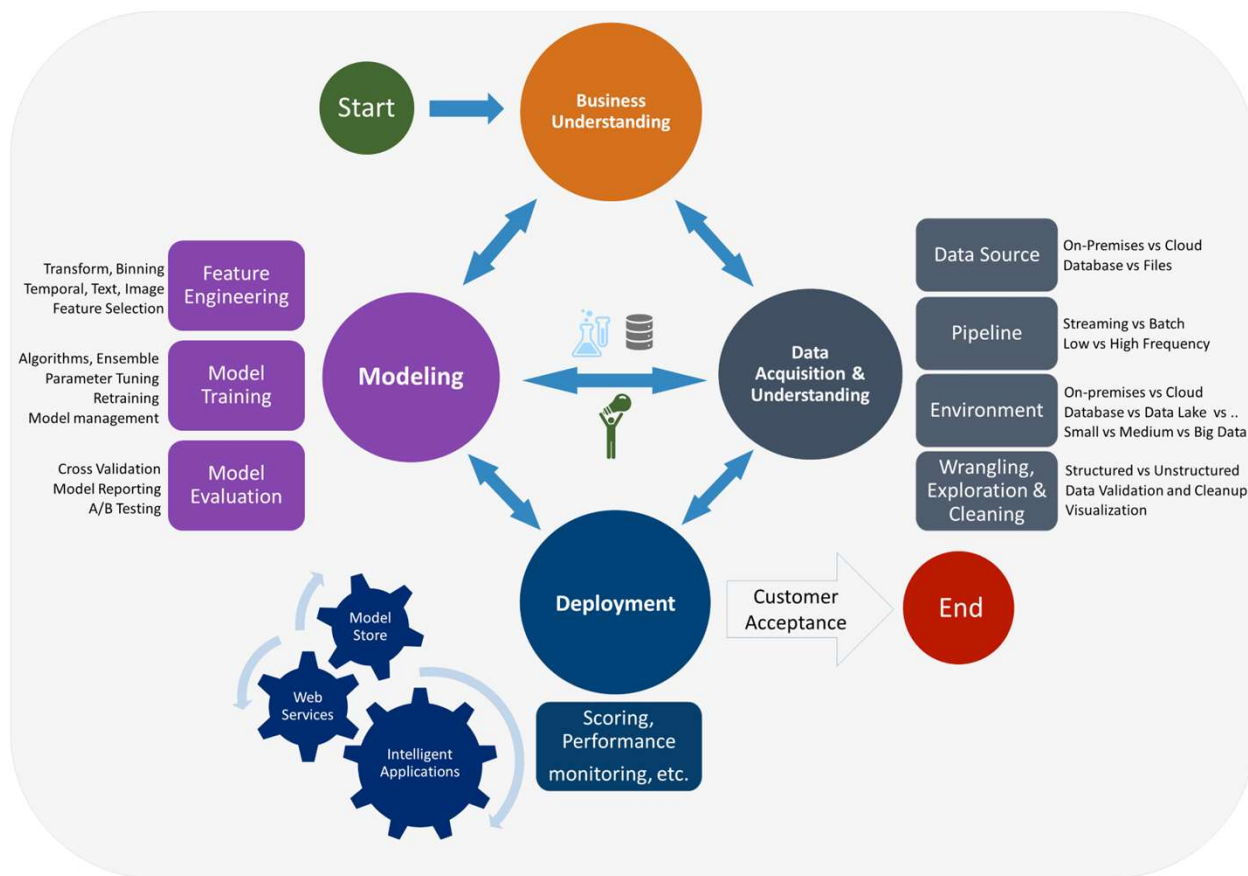
Figura 1: Visión general de los pasos que componen el proceso KDD

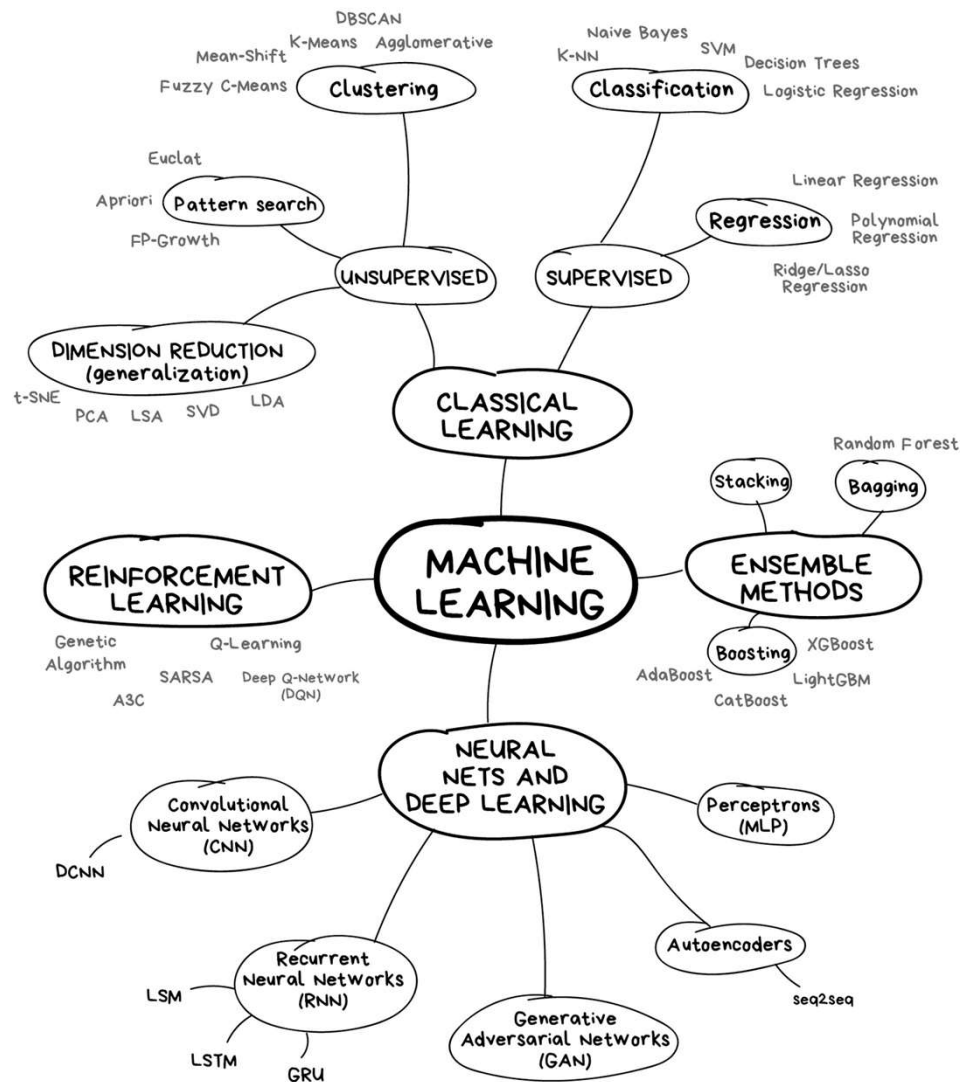
CRISP-DM (Cross-Industry Standard Process for Data Mining)



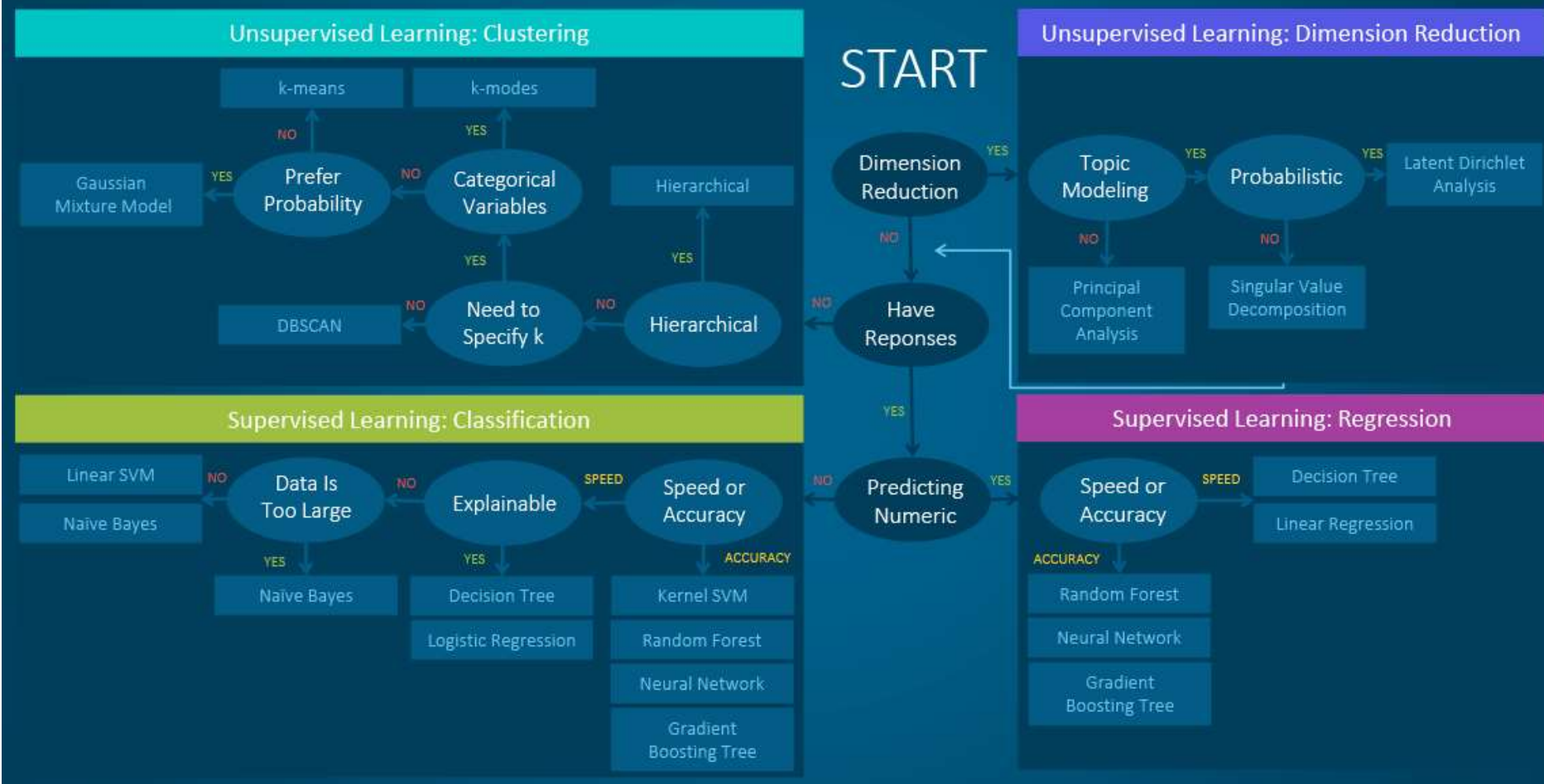
t-DSP (Team Data Science Process)

Data Science Lifecycle

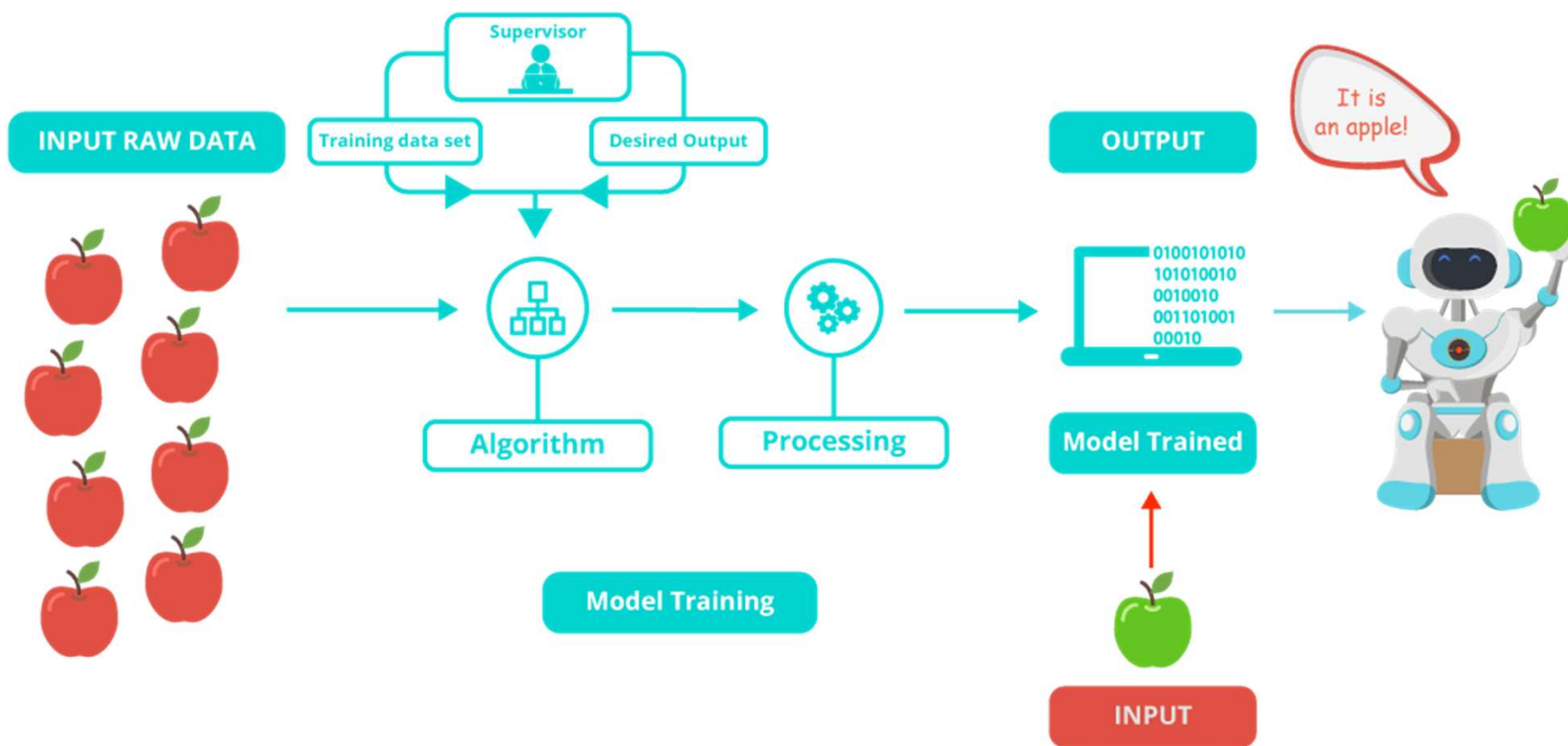




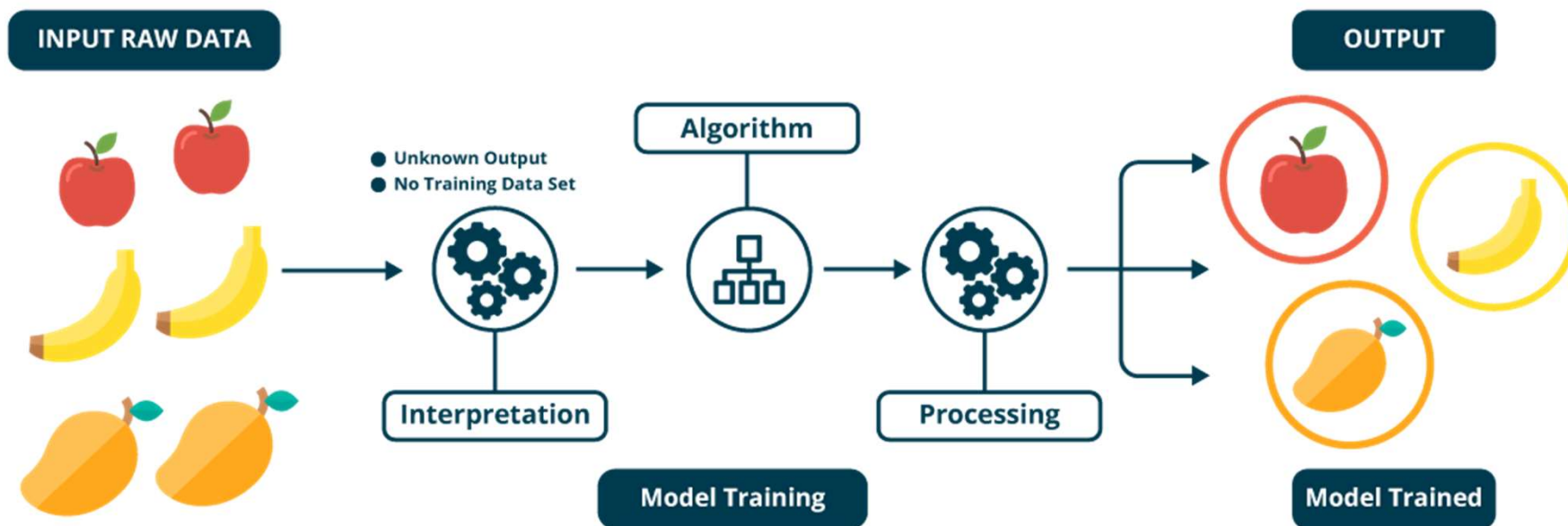
Machine Learning Algorithms Cheat Sheet



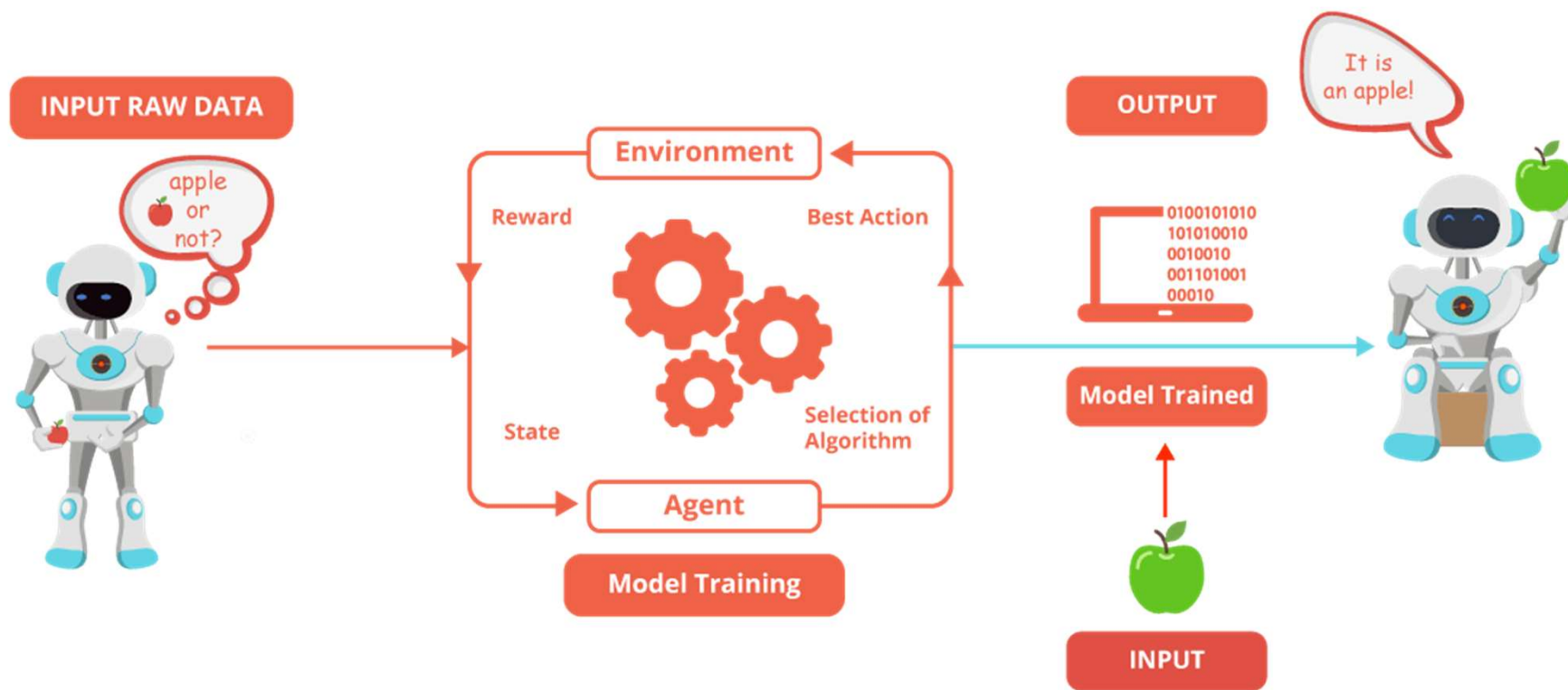
Aprendizaje Supervisado



Aprendizaje No supervisado



Aprendizaje por Reforzamiento



Entrenamiento de un modelo de aprendizaje de máquina

Color (nm)	Alcohol %	Beer or Wine?
610	5	Beer
599	13	Wine
693	14	Wine

$$y = m * x + b$$

OUTPUT SLOPE INPUT Y-INTERCEPT

WEIGHTS =

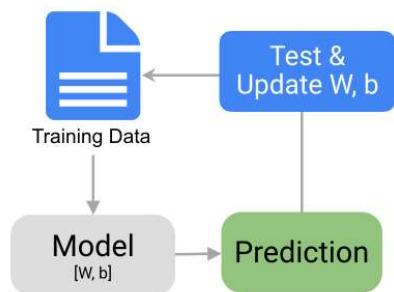
$$\begin{bmatrix} m_{1,1} & m_{1,2} \\ m_{2,1} & m_{2,2} \\ m_{3,1} & m_{3,2} \end{bmatrix}$$

BIASES =

$$\begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \\ b_{3,1} & b_{3,2} \end{bmatrix}$$

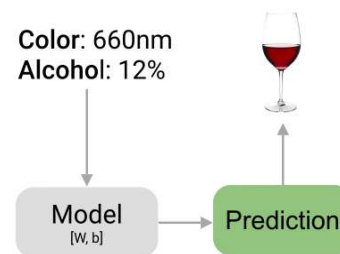
1) Conjunto de Datos
(Obtención, Preprocesamiento, Estandarización, Transformación, etc)

2) Selección de un Modelo a partir del objetivo deseado y de los datos de entrada



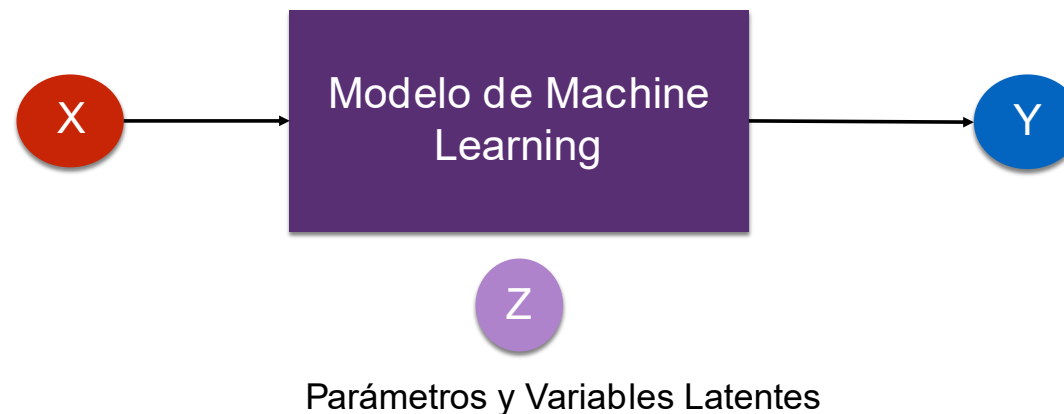
3) Ajuste del modelo (parámetros)

A través de un algoritmo de aprendizaje/optimización



4) Evaluación y puesta en práctica

Variables en los modelos de Machine Learning



X, Y, Z son Variables Aleatorias con distribuciones de Probabilidad

$$P(X,Y,Z)$$

VARIABLES EN LOS MODELOS DE MACHINE LEARNING

A partir de la Probabilidad Conjunta

$$P(X, Y, Z)$$

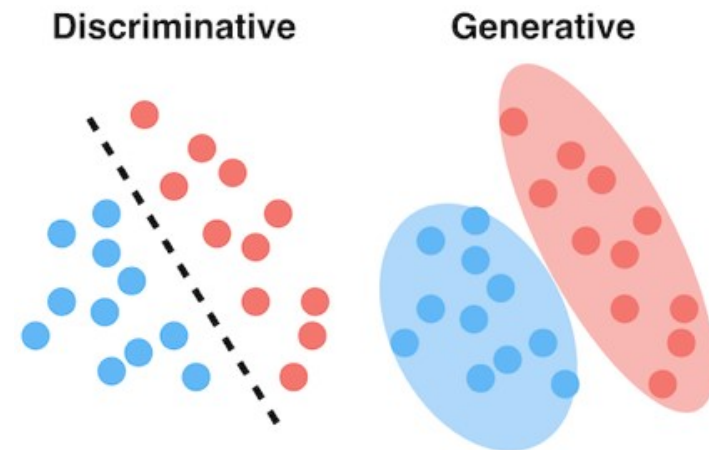
Se puede estimar

Modelo Discriminativo

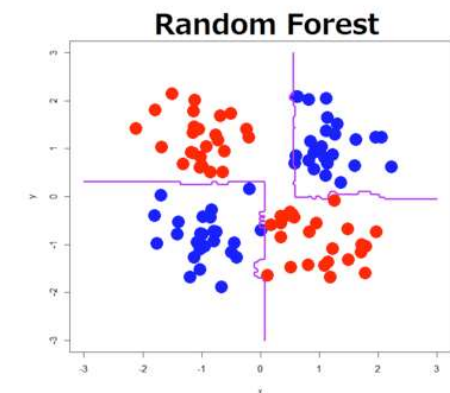
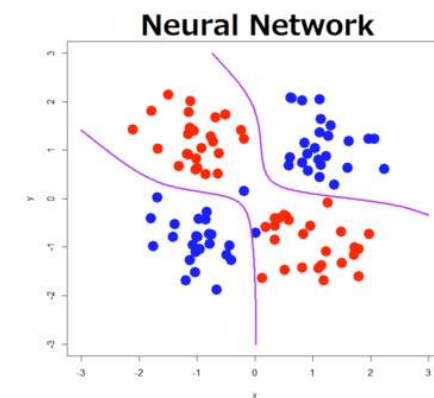
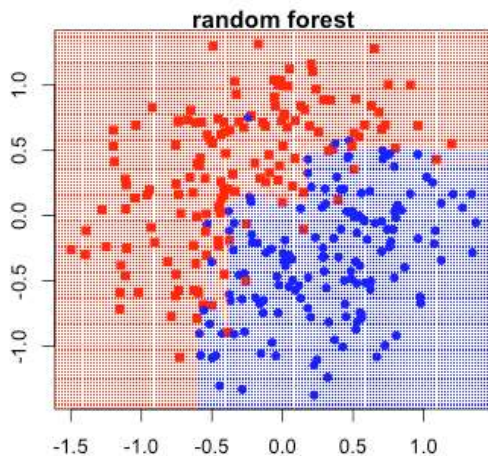
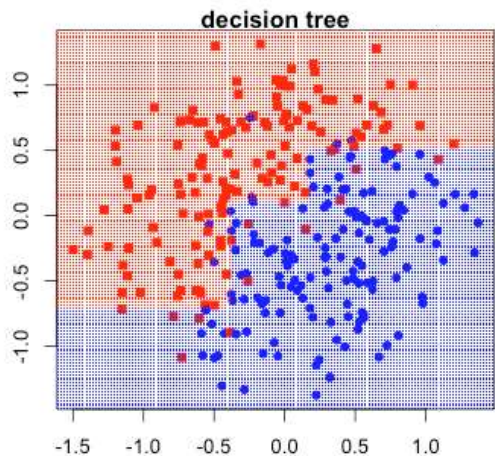
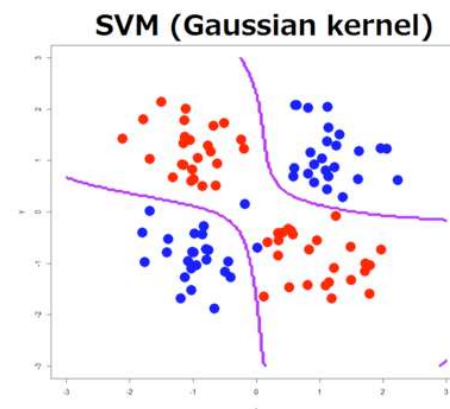
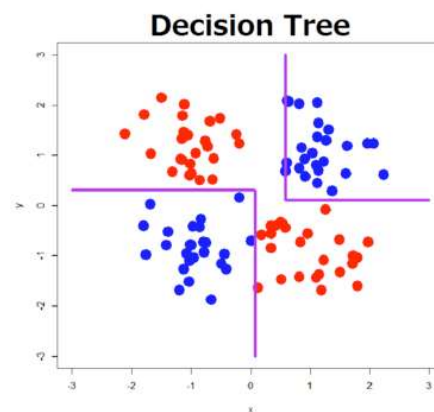
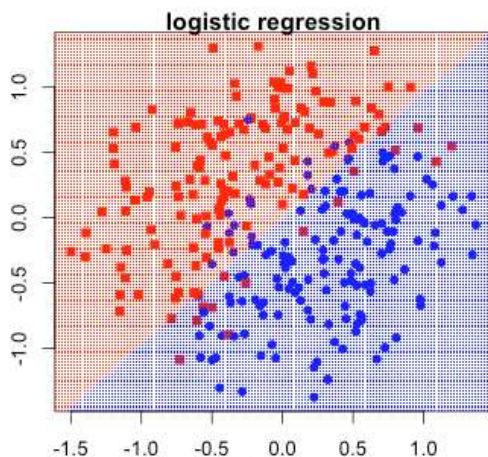
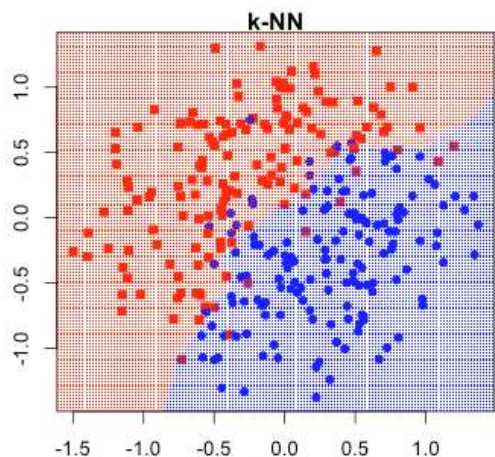
$$P(Y|X, Z) =$$

Modelo Generativo

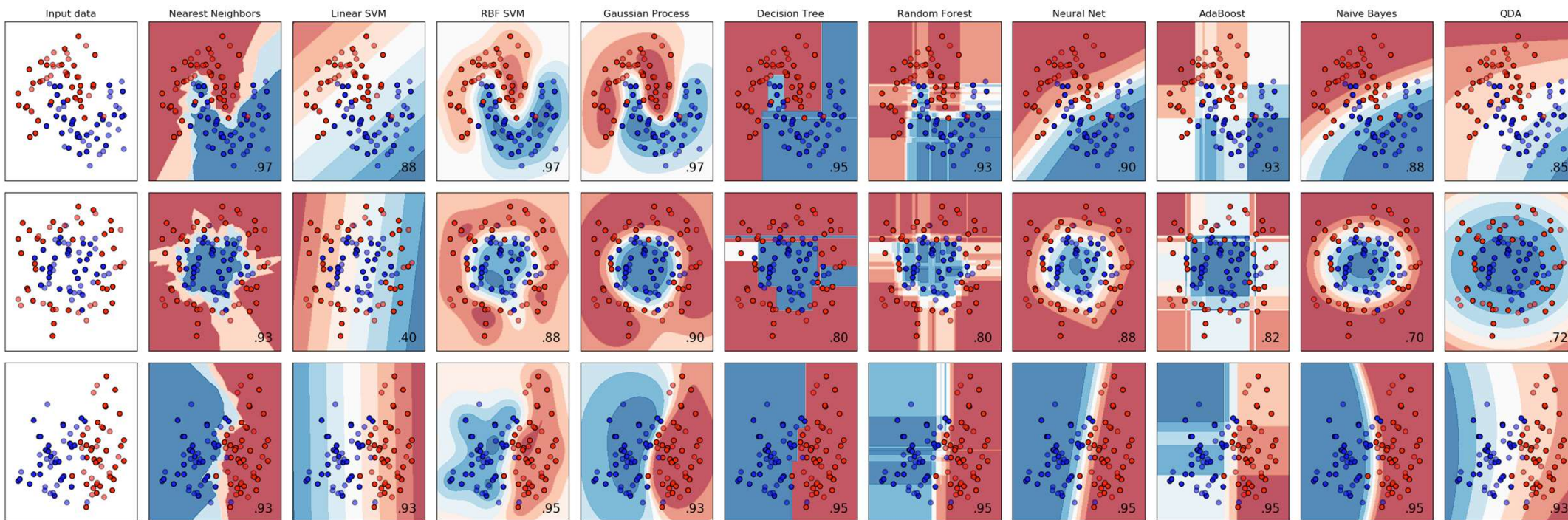
$$P(X|Y, Z) =$$



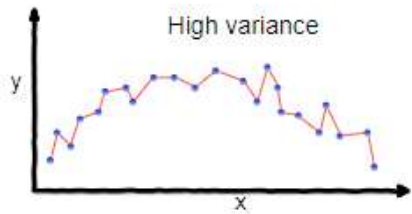
Frnteras de decisi3n



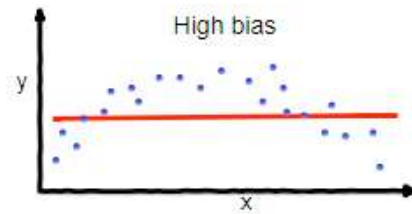
Clustering



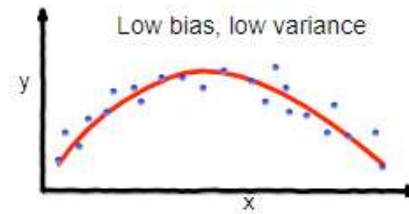
Sesgo y varianza de un modelo (Bias and Variance)



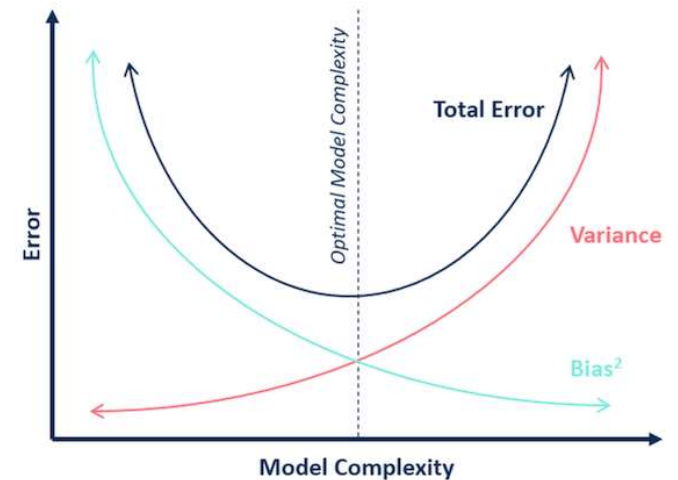
overfitting



underfitting



Good balance



Conjunto de datos

[Obesity Levels & Life Style | Kaggle](https://www.kaggle.com/code/mpwolke/obesity-levels-life-style)

<https://www.kaggle.com/code/mpwolke/obesity-levels-life-style>

Dataset for estimation of obesity levels based on eating habits and physical condition in individuals from Colombia, Peru and Mexico

Authors: Fabio MendozaPalechor;; Alexis de la HozManotas <https://doi.org/10.1016/j.dib.2019.104344>

This paper presents data for the estimation of obesity levels in individuals from the countries of Mexico, Peru and Colombia, based on their eating habits and physical condition. The data contains 17 attributes and 2111 records, the records are labeled with the class variable NObesity (Obesity Level), that allows classification of the data using the values of Insufficient Weight, Normal Weight, Overweight Level I, Overweight Level II, Obesity Type I, Obesity Type II and Obesity Type III.

Attributes related with eating habits are:

- Frequent consumption of high caloric food (FAVC),
- Frequency of consumption of vegetables (FCVC),
- Number of main meals (NCP),
- Consumption of food between meals (CAEC),
- Consumption of water daily (CH20),
- and Consumption of alcohol (CALC).

Attributes related with the physical condition are:

- Calories consumption monitoring (SCC),
- Physical activity frequency (FAF),
- Time using technology devices (TUE),
- Transportation used (MTRANS),

Other variables were:

- Gender, Age, Height and Weight.

Gender	Age	Height	Weight	family_history	FAVC	FCVC	NCP	CAEC	SMOKE	CH20	SCC	FAF	TUE	CALC	MTRANS	NObesidad	
Female	21	1.62	64	yes	no	2	3	Sometimes	no	2	no		0	1	no	Public_Transportation	Normal_Weight
Female	21	1.52	56	yes	no	3	3	Sometimes	yes	3	yes		3	0	Sometimes	Public_Transportation	Normal_Weight
Male	23	1.8	77	yes	no	2	3	Sometimes	no	2	no		2	1	Frequently	Public_Transportation	Normal_Weight
Male	27	1.8	87	no	no	3	3	Sometimes	no	2	no		2	0	Frequently	Walking	Overweight_Level_I
Male	22	1.78	89.8	no	no	2	1	Sometimes	no	2	no		0	0	Sometimes	Public_Transportation	Overweight_Level_II
Male	29	1.62	53	no	yes	2	3	Sometimes	no	2	no		0	0	Sometimes	Automobile	Normal_Weight
Female	23	1.5	55	yes	yes	3	3	Sometimes	no	2	no		1	0	Sometimes	Motorbike	Normal_Weight
Male	22	1.64	53	no	no	2	3	Sometimes	no	2	no		3	0	Sometimes	Public_Transportation	Normal_Weight
Male	24	1.78	64	yes	yes	3	3	Sometimes	no	2	no		1	1	Frequently	Public_Transportation	Normal_Weight
Male	22	1.72	68	yes	yes	2	3	Sometimes	no	2	no		1	1	no	Public_Transportation	Normal_Weight
Male	26	1.85	105	yes	yes	3	3	Frequently	no	3	no		2	2	Sometimes	Public_Transportation	Obesity_Type_I
Female	21	1.72	80	yes	yes	2	3	Frequently	no	2	yes		2	1	Sometimes	Public_Transportation	Overweight_Level_II
Male	22	1.65	56	no	no	3	3	Sometimes	no	3	no		2	0	Sometimes	Public_Transportation	Normal_Weight

¡Gracias!

¿Alguna pregunta?

julopez@uv.mx

www.uv.mx/fei



www.uv.mx/linae



Contáctanos

