

[www.mtydigitalhub.com](http://www.mtydigitalhub.com)

**digital**  
hub **MTY**  
LEARN · WORK · CONNECT



# Bienvenido Data Translator

## Objetivos del Módulo:



- El estudiante tendrá una reconexión con sus compañeros
- El estudiante tendrá un repaso de el algoritmo knn
- El estudiante implementara knn usando la biblioteca sklearn

# Vamos a re-conocernos

[www.mtydigitalhub.com](http://www.mtydigitalhub.com)

**digital**  
hub **MTY**  
LEARN · WORK · CONNECT



De manera anónima hagan una carta en Word, en donde se hablarían en el espejo, ¿Qué se dirían?, ¿Qué palabras de aliento creen que harían una diferencia en su día?

Al finalizar me lo envían por un mensaje de zoom directo, esto quiere decir que podré leerlo por ustedes.

Es probable que más de una persona en el grupo quiera escuchar esto, o **¡LO NECESITE!**

**Todos estamos pasando por momentos difíciles**



**Cuaderno (de cuadro de preferencia)**

**Lapiz**

**Goma**

**\* Color rojo y azul**

**Calculadora**

**¡Café!**



**¡En sus marcas! ¡Listos! K-NN**

El siguiente es un ejemplo simple en el que clasificaremos el punto (4, 3) para saber si es de color azul o rojo. Tenemos los siguientes puntos con sus respectivos colores:

Rojos	Azules
(1, 1)	(1, 4)
(1, 2)	(1, 7)
(2, 4)	(2, 2)
(2, 5)	(2, 7)
(3, 1)	(2, 8)
(3, 3)	(3, 4)
(4, 2)	(3, 6)
(4, 5)	(4, 4)
(4, 7)	(4, 6)
(5, 2)	(5, 4)
(5, 3)	(5, 5)
(6, 3)	(5, 8)
(7, 1)	(6, 4)
(7, 4)	(6, 5)
	(7, 5)

El siguiente es un ejemplo simple en el que clasificaremos el punto (4, 3) para saber si es de color azul o rojo. Tenemos los siguientes puntos con sus respectivos colores:

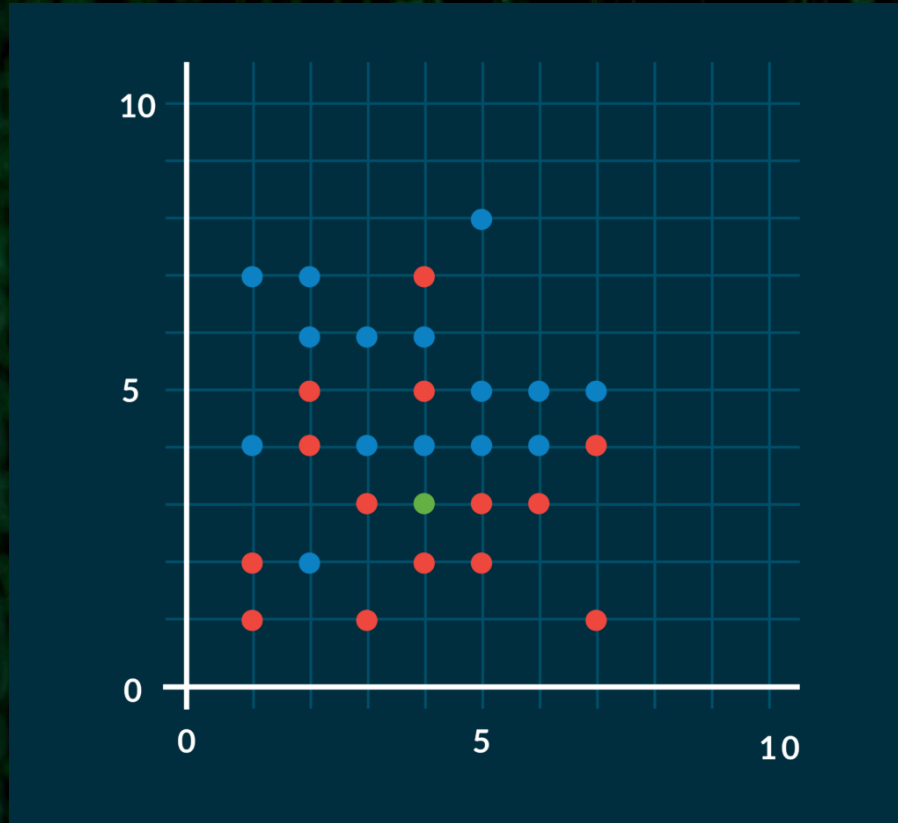
Rojos		Azules	
(1, 1)	1	(1, 4)	1
(1, 2)	2	(1, 7)	2
(2, 4)	3	(2, 2)	3
(2, 5)	4	(2, 7)	4
(3, 1)	5	(2, 8)	5
(3, 3)	6	(3, 4)	6
(4, 2)	7	(3, 6)	7
(4, 5)	8	(4, 4)	8
(4, 7)	9	(4, 6)	9
(5, 2)	10	(5, 4)	10
(5, 3)	11	(5, 5)	11
(6, 3)	12	(5, 8)	12
(7, 1)	13	(6, 4)	13
(7, 4)	14	(6, 5)	14
		(7, 5)	14

El siguiente es un ejemplo simple en el que clasificaremos el punto (4, 3) para saber si es de color azul o rojo. Tenemos los siguientes puntos con sus respectivos colores:

Rojos		Azules	
(1, 1)	1	(1, 4)	1
(1, 2)	2	(1, 7)	2

Tener 13 datos de una categoría y 14 datos de otra categoría, se le conoce como “data balanceada”, lo cual solo lo vemos en el mundo ideal.

(5, 2)	9	(5, 4)	9
(5, 3)	10	(5, 5)	10
(6, 3)	11	(5, 8)	11
(7, 1)	12	(6, 4)	12
(7, 4)	13	(6, 5)	13
		(7, 5)	14





Usando la distancia euclidiana, podemos calcular la distancia entre el punto (4, 3) y los demás puntos. Para este ejemplo seleccionaremos  $k = 7$  vecinos más cercanos.

**¡Si! nuestros 27 puntos vs el punto (4,3)**

Tip: organícense como grupo que son para dividirse los puntos y terminar antes, colaboremos todos y utilicemos el chat de zoom!



Rojos	Distancia	Azules	Distancia
(1, 1)	3.605551275	(1, 4)	3.16227766
(1, 2)	3.16227766	(1, 7)	5
(2, 4)	2.236067977	(2, 2)	2.236067977
(2, 5)	2.828427125	(2, 7)	4.472135955
(3, 1)	2.236067977	(2, 8)	5.385164807
(3, 3)	1	(3, 4)	1.414213562
(4, 2)	1	(3, 6)	3.16227766
(4, 5)	2	(4, 4)	1
(4, 7)	4	(4, 6)	3
(5, 2)	1.414213562	(5, 4)	1.414213562
(5, 3)	1	(5, 5)	2.236067977
(6, 3)	2	(5, 8)	5.099019514
(7, 1)	3.605551275	(6, 4)	2.236067977
(7, 4)	3.16227766	(6, 5)	2.828427125
		(7, 5)	3.605551275

Como se puede ver en la tabla los puntos (3, 3), (3, 4), (4, 2), (4, 4), (5, 2), (5, 3) y (5, 4) son los puntos con menor distancia al punto (4, 3), de los cuales 4 son rojos y 3 son azules. Por lo tanto, clasificamos el punto (4, 3) como rojo.

Rojos	Distancia	Azules	Distancia
(1, 1)	3.605551275	(1, 4)	3.16227766
(1, 2)	3.16227766	(1, 7)	5
(2, 4)	2.236067977	(2, 2)	2.236067977
(2, 5)	2.828427125	(2, 7)	4.472135955
(3, 1)	2.236067977	(2, 8)	5.385164807
(3, 3)	1	(3, 4)	1.414213562
(4, 2)	1	(3, 6)	3.16227766
(4, 5)	2	(4, 4)	1
(4, 7)	4	(4, 6)	3
(5, 2)	1.414213562	(5, 4)	1.414213562
(5, 3)	1	(5, 5)	2.236067977
(6, 3)	2	(5, 8)	5.099019514
(7, 1)	3.605551275	(6, 4)	2.236067977
(7, 4)	3.16227766	(6, 5)	2.828427125
		(7, 5)	3.605551275

Rojos	Distancia	Azules	Distancia
(1, 1)	3.605551275	(1, 4)	3.16227766
(1, 2)	3.16227766	(1, 7)	5
(2, 4)	2.236067977	(2, 2)	2.236067977
(2, 5)	2.828427125	(2, 7)	4.472135955
(3, 1)	2.236067977	(2, 8)	5.385164807
(3, 3)	1	(3, 4)	1.414213562
(4, 2)	1	(3, 6)	3.16227766
(4, 5)	2	(4, 4)	1
(4, 7)	4	(4, 6)	3
(5, 2)	1.414213562	(5, 4)	1.414213562
(5, 3)	1	(5, 5)	2.236067977
(6, 3)	2	(5, 8)	5.099019514
(7, 1)	3.605551275	(6, 4)	2.236067977
(7, 4)	3.16227766	(6, 5)	2.828427125
		(7, 5)	3.605551275

$$4 \text{ Rojos} = 4/7 = 0.57$$

$$3 \text{ Azules} = 3/7 = 0.42$$



Por lo tanto podemos concluir que el punto (4,3) es un punto:

**ROJO**

Rojos	Distancia	Azules	Distancia
(1, 1)	3.605551275	(1, 4)	3.16227766
(1, 2)	3.16227766	(1, 7)	5
(2, 4)	2.236067977	(2, 2)	2.236067977
(2, 5)	2.828427125	(2, 7)	4.472135955
(3, 1)	2.236067977	(2, 8)	5.385164807
(3, 3)	1	(3, 4)	1.414213562
(4, 2)	1	(3, 6)	3.16227766
(4, 5)	2	(4, 4)	1
(4, 7)	4	(4, 6)	3
(5, 2)	1.414213562	(5, 4)	1.414213562
(5, 3)	1	(5, 5)	2.236067977
(6, 3)	2	(5, 8)	5.099019514
(7, 1)	3.605551275	(6, 4)	2.236067977
(7, 4)	3.16227766	(6, 5)	2.828427125
		(7, 5)	3.605551275

$$4 \text{ Rojos} = 4/7 = 0.57$$

$$3 \text{ Azules} = 3/7 = 0.42$$

[www.mtydigitalhub.com](http://www.mtydigitalhub.com)

**digital**  
hub **MTY**  
LEARN · WORK · CONNECT



# Sk-learn with knn

## Pasos a seguir al momento demodelar

- 1.- Bussines Objctive
- 2.- Data requirement
- 3.-Data Collection
- 4.- EDA --- ¿Qué significa?
- 5.- Modeling
- 6.-Evaluation
- 7.-Deployment
- 8.-Monitoring

## Pasos a seguir al momento demodelar

### 1.- Bussines Objective

El objetivo es tratar de identificar el momento en que ocurre una falla en la línea de producción.



Pasos a seguir al  
momento demodelar

## 2.- Data requirement

¿Qué datos necesito para lograr mi objetivo?

Pasos a seguir al  
momento demodelar

### 3.-Data Collection

- ¿Cómo puedo tener los datos necesito para lograr mi objetivo?
- ¿Con que granularidad los necesito?
- ¿Cuáles datos son fundamentales, y cuales son de “preferencia”?
- ¿Con cuanto tiempo de antelación los necesito? (Histórico)

## Objetivos del Módulo:



Pasos a seguir al  
momento demodelar

## 4.- EDA (Exploratory Data Analysis)

### Análisis exploratorio de datos

¿Qué me puedo preguntar?



## Objetivos del Módulo:



Pasos a seguir al  
momento demodelar

## 5.- Modeling

Hasta este punto apenas vamos a hacer el modelo.  
¿List@?

## Objetivos del Módulo:



Pasos a seguir al  
momento demodelar

6.-Evaluation

7.-Deployment

8.-Monitoring

Lo verán más adelante



[www.mtydigitalhub.com](http://www.mtydigitalhub.com)

**digital**  
hub **MTY**  
LEARN · WORK · CONNECT



# ¡Gracias!

---

# Data Translator