



BOURBAKI

COLEGIO DE MATEMÁTICAS

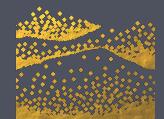
alfonso@escuela-bourbaki.com



A N A I S A B E L A S C E N C I O P E D R A Z A



Ana Isabel Ascencio Pedraza es Científica de Datos con más de 20 años de experiencia en análisis de datos para la toma de decisiones. Estudió Ingeniería Electromecánica en la Universidad Iberoamericana León, Métodos Estadísticos en el Centro de Investigaciones en Matemáticas (CIMAT) y Ciencia de Datos en el Centro de Investigación e Innovación en Tecnologías de la Información (INFOTEC). Actualmente es consultora freelance en Ciencia de Datos y Analítica Avanzada.



¡Les damos la bienvenida!

EL PERCEPTRÓN

- Caso de uso (una hora)
- Un vistazo a Python (una hora)
- Vectorización, hiperplanos y el perceptrón (dos horas)
- Cáncer de esófago (dos horas)
- Dudas y KC (una hora)
- Overfitting (una hora)

Caso de uso de esta semana:

C L A S I F I C A C I Ó N
D E T E J I D O S
C A N C E R Í G E N O S





Mauna Kea Technologies

Mauna Kea Technologies es una compañía global de dispositivos médicos enfocada en eliminar incertidumbres relacionadas con el diagnóstico y tratamiento del cáncer y otras enfermedades gracias a visualización microscópica **en tiempo real**. El producto estrella de la Compañía, Cellvizio®, ha sido autorizado en más de 40 países.

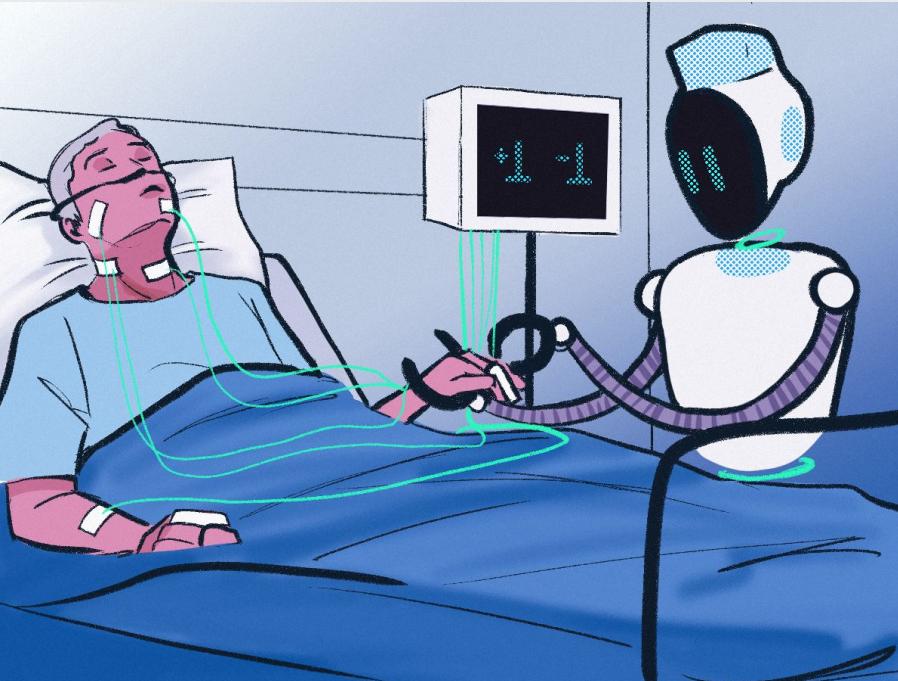


Cáncer de esófago

El cáncer del esófago es una enfermedad común en los pacientes que presentan la enfermedad por reflujo gastroesofágico la cual es muy común en adultos, por ejemplo el 7% de la población en USA la tiene.



Detección costosa



Tradicionalmente es necesario que un médico patólogo analiza una muestra física del esófago y a partir de ahí pueda diagnosticar si el paciente tiene cáncer de esófago.



Mini microscopio

Mauna Kea desarrolló un microscopio y una tecnología para capturar imágenes microscópicas del esófago utilizando un endoscopio:

Cellvizio®



Utilizar Machine Learning



El objetivo de este proyecto es utilizar un modelo de machine learning que le permita al médico concentrarse en algunas imágenes particularmente importantes. Facilitar el trabajo del médico, no reemplazarlo.

En términos técnicos estamos buscando un modelo de clasificación binaria de imágenes.





Cellvizio® enables physicians to **monitor** the progression of disease over time, **assess** point-in-time reactions as they happen in real time, **classify** indeterminate areas of concern, and **guide** surgical interventions.

Confocal Miniprobes™

for a wide range of medical specialties

The anatomical diagram illustrates the locations where the various Cellvizio miniprobes can be used. Blue lines connect the probe names to their corresponding anatomical sites: GastroFlex™ UHD to the esophagus; AlveoFlex™ to the lungs; CholangioFlex™ to the bile ducts; AQ-Flex™ 19 to the trachea; CelioFlex™ UHD 5 to the liver; ColoFlex™ UHD to the colon; UroFlex™ B to the bladder; CystoFlex™ F to the prostate; and CystoFlex™ UHD R to the rectum.

	Compatible operating channel	Length	Maximum # of uses	Field of view	Resolution	Working Distance
GastroFlex™ UHD	≥ 2.8 mm	3m	20	Ø 240 µm	1 µm	60 +/- 15 µm
AlveoFlex™	≥ 1.9 mm	3m	20	Ø 600 µm	3.5 µm	>0 µm
CholangioFlex™	≥ 1.0 mm	4m	10	Ø 325 µm	3.5 µm	55 +/- 15 µm
AQ-Flex™ 19	≥ 0.91 mm (19 Fr)	4m	10	Ø 325 µm	3.5 µm	55 +/- 15 µm
CelioFlex™ UHD 5	≥ 5 mm	3m	20	Ø 240 µm	1 µm	60 +/- 15 µm
ColoFlex™ UHD	≥ 2.8 mm	4m	20	Ø 240 µm	1 µm	60 +/- 15 µm
UroFlex™ B	≥ 1 mm (3 Fr)	3m	10	Ø 325 µm	3.5 µm	55 +/- 15 µm
CystoFlex™ F	≥ 1 mm (3 Fr)	2m	20	Ø 325 µm	3.5 µm	55 +/- 15 µm
CystoFlex™ UHD R	≥ 2.8 mm (8.4 Fr)	2m	20	Ø 240 µm	1 µm	60 +/- 15 µm

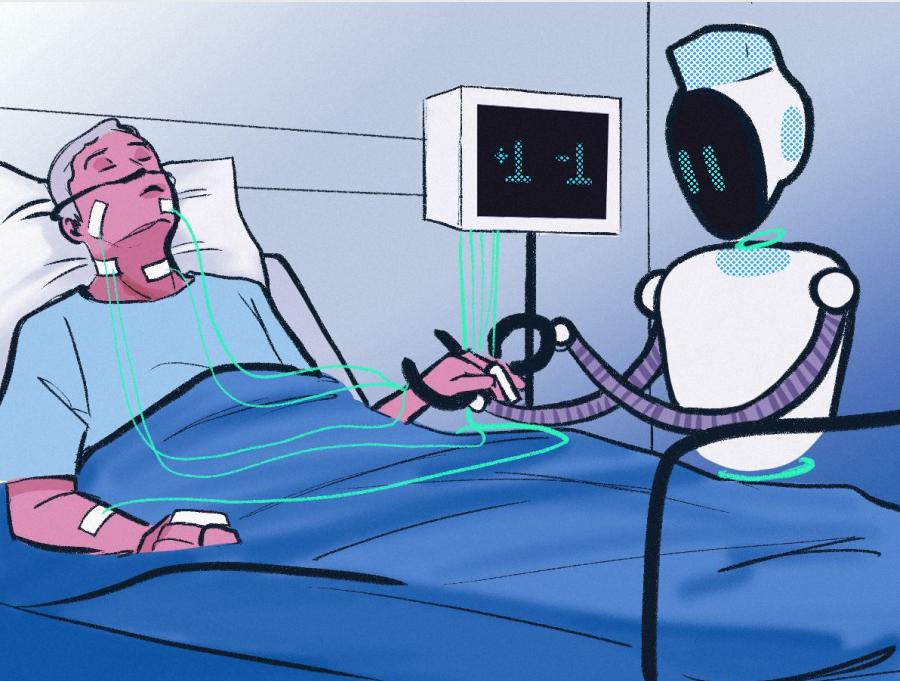
Compatible with any standard endoscope and reprocessing method

Cellvizio

Aprobado por la Food & Drugs Administration (FDA) es utilizado en diversos países y se utiliza en conjunto con profesionales de la salud para hacer diagnósticos más precisos.



Sutilezas

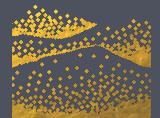


- ¿Base de datos?
- Imágenes inestables.
- Errores del segundo tipo.



S U P E R V I S A D O

Aprendizaje



(x, y)



Altamente costoso



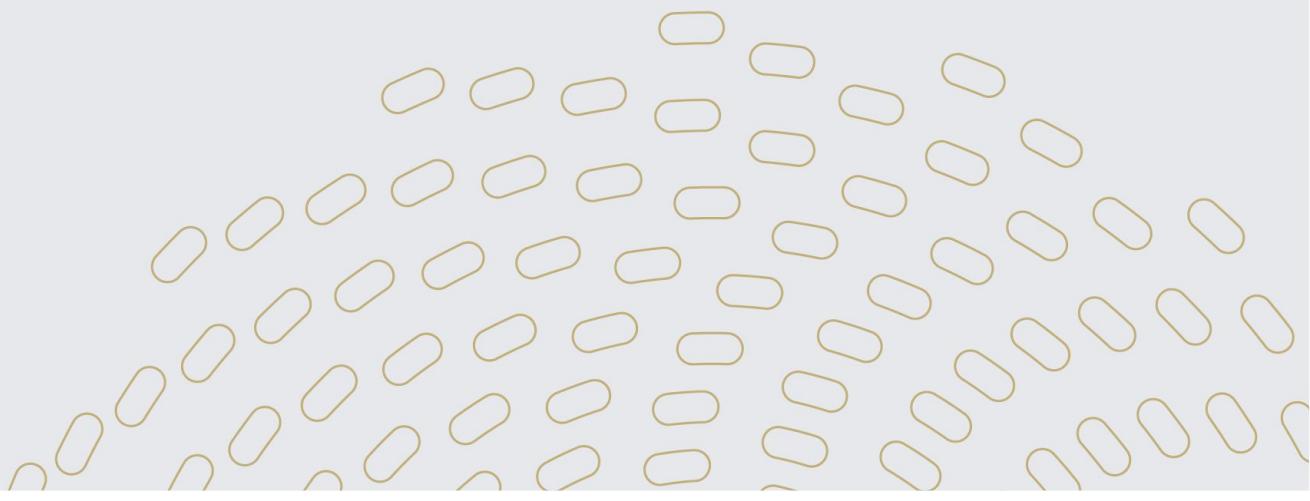
¡No deseamos
clasificar a la
base de datos!



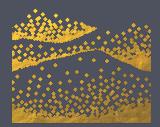
¿P A R A Q U É
S I R V E?



¿O T R A S
A P L I C A C I O N E S?



DOS ETAPAS



Etapa uno

Aprendizaje

Algoritmo

Etapa dos

Modelo

Predictión

Nuestro algoritmo
=
PERCEPTRÓN



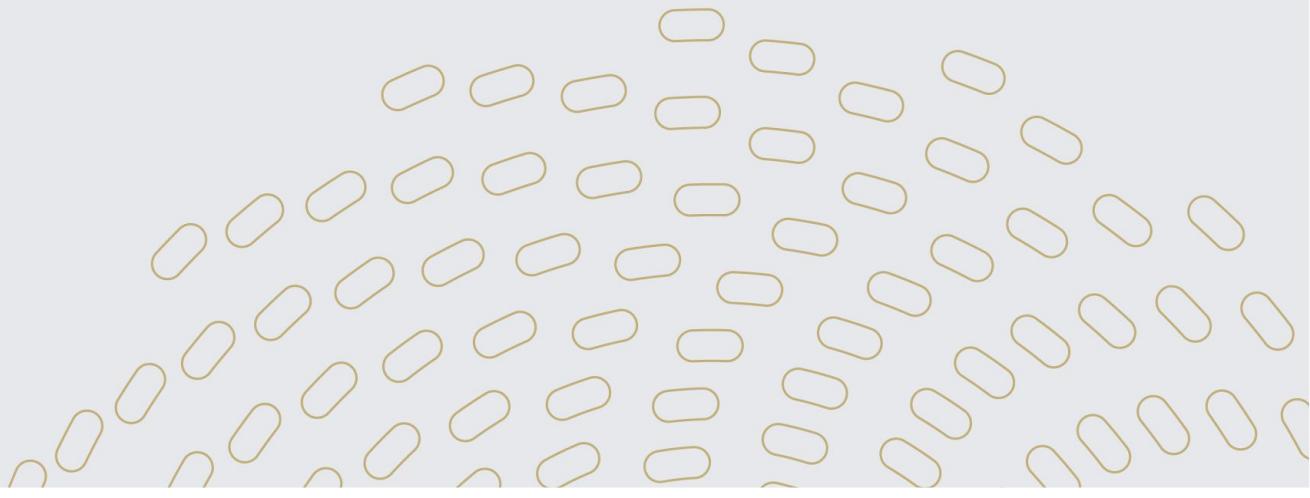
Nuestro modelo

=

H I P E R P L A N O



¿ CÓMO FUNCIONAN LOS HIPERPLANOS ?



Ponderar

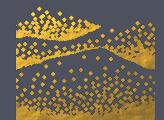
C A R A C T E R Í S T I C A S



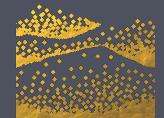
Caracterizar

=

V E C T O R I Z A R



Vectorizar
el
D A T A S E T





¿ V E C T O R E S ?

¿Es necesario?

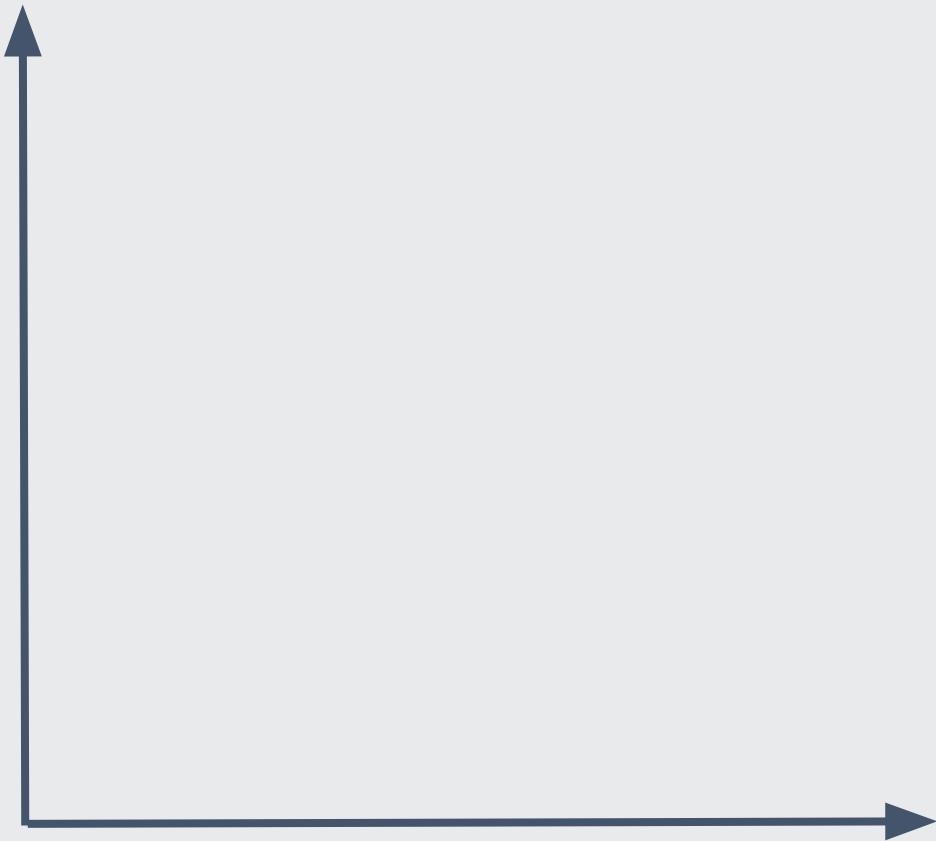


$$x = (x_1, x_2, \dots, x_d)$$



\mathbb{R}^d 

\mathbb{R}^2 



$$d = 260^2$$



$$x = (1, 1, \dots, 1)$$



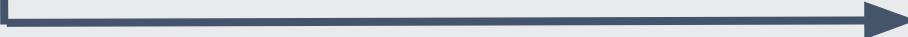
$$x = (0, 0, \dots, 0)$$




$$(x_1, x_2)$$


$$y \in \{-1,+1\}$$




$$(x_1, x_2)$$


$$\beta = (\beta_1,\beta_2,\ldots,\beta_d)$$



$\beta = ?$



$$\langle \beta, x \rangle = \sum_{i \leq d} (\beta_i x_i) = \beta_1 x_1 + \dots + \beta_d x_d$$



¿Cómo evaluar un modelo de
M A C H I N E
L E A R N I N G ?



Dividir las

5,000 IMÁGENES

