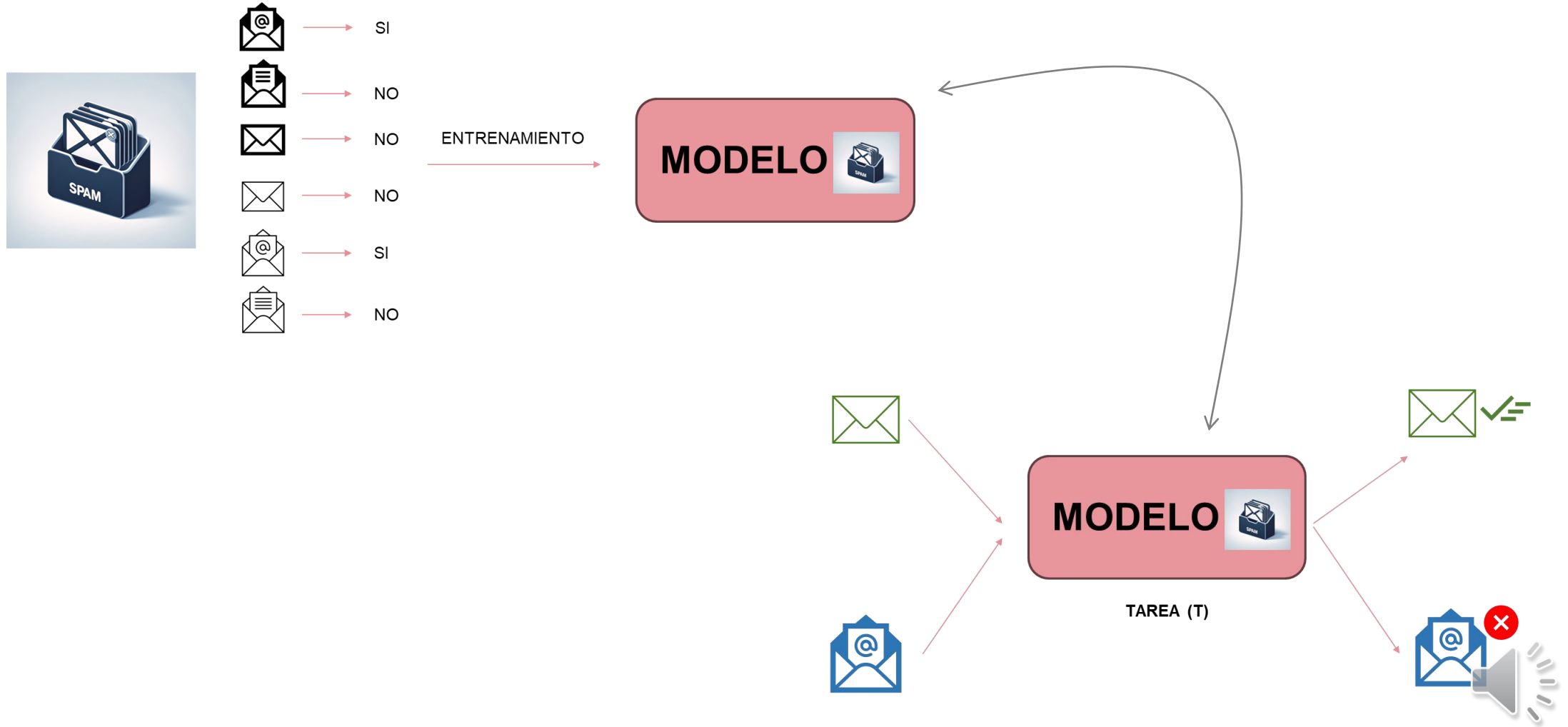




Introducción a Machine Learning Modelos







Modelos: ¿Qué es “físicamente”?

- Una pieza de código.
- En nuestro caso, y en general, un objeto Python de una clase que depende del tipo de algoritmo utilizado
- Con métodos para entrenar/ajustar el modelo (que tendrán el training set como input)
- Con métodos para predecir/darnos una respuesta ante datos nuevos e instancias nuevas (al que pasaremos un dataframe en general)
- Además el modelo tendrá sus métodos para evaluar su rendimiento, aunque a veces tendremos que tirar de funciones externas
- En general, vamos a ver tantos tipos de clase como tipos de algoritmos aplicables.



Modelos: ¿Qué es conceptualmente?

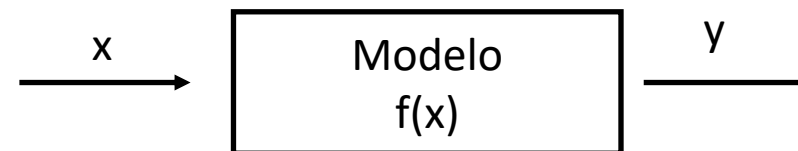
Un modelo es un conjunto de fórmulas matemáticas que expresan relaciones y patrones entre variables, utilizados para estudiar comportamientos de sistemas complejos ante situaciones difíciles de observar en la realidad.

Es una manera de traducir a la matemática relaciones entre datos de la vida real.

Variables que me digan la predisposición que tienen los pacientes a contraer cáncer de pulmón

Variables que me indiquen si un cliente es buen o mal pagador para un banco.

Qué producto estaría dispuesto a comprar un usuario X de Facebook, en función de sus gustos y búsquedas.



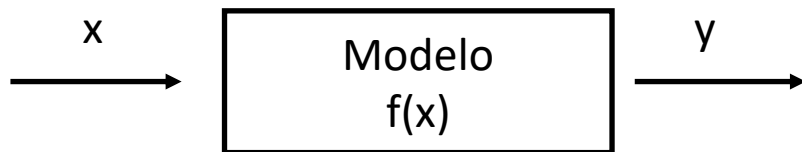
Con un modelo puedo predecir cierto output, como por ejemplo si el paciente contrae o no cáncer de pulmón, a partir de un conjunto de inputs, que podrían ser edad, fumador, enfermedades previas...



Modelos: Target (y) y Features (x)

Target (Y)

Se trata de la variable objetivo, la variable a predecir, el output del modelo. Por norma general es una única variable, ya sea numérica o categórica

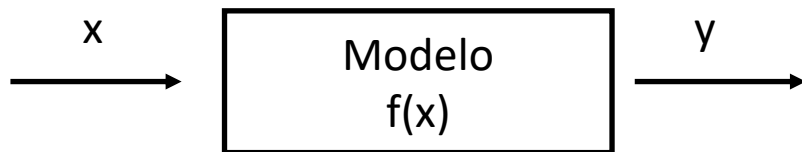


Modelos: Target (y) y Features (x)

Target (Y)

Se trata de la variable objetivo, la variable a predecir, el output del modelo. Por norma general es una única variable, ya sea numérica o categórica

SPAM: El target es el tipo de correo: Spam o no. Variable categórica con posibles valores: "Spam"/"No Spam", o "True/False", o "1/0"



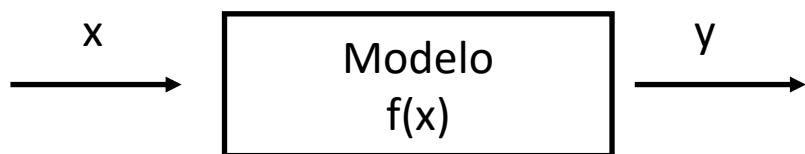
Modelos: Target (y) y Features (x)

Target (Y)

Se trata de la variable objetivo, la variable a predecir, el output del modelo. Por norma general es una única variable, ya sea numérica o categórica

SPAM: El target es el tipo de correo: Spam o no. Variable categórica con posibles valores: "Spam"/"No Spam", o "True/False", o "1/0"

CASAS: El target es el precio. Variable numérica continua. Problema de regresión.



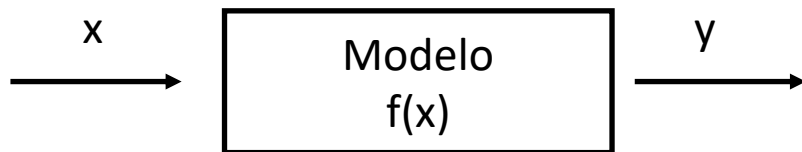
Modelos: Target (y) y Features (x)

Target (Y)

Se trata de la variable objetivo, la variable a predecir, el output del modelo. Por norma general es una única variable, ya sea numérica o categórica

Features (X)

El resto de las variables de un dataset utilizadas como input del modelo (para aprender y para predecir)



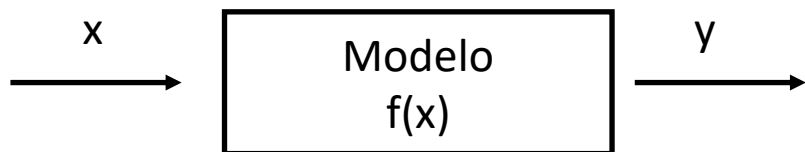
Modelos: Target (y) y Features (x)





Target (Y)

Se trata de la variable objetivo, la variable a predecir, el output del modelo. Por norma general es una única variable, ya sea numérica o categórica

Features (X)

El resto de las variables de un dataset utilizadas como input del modelo (para aprender y para predecir)



	X			Y
	Fumador	Edad	¿Ha tenido cáncer?	
	Si	57	Si	
	No	32	No	
	Si	39	No	
	Si	60	No	



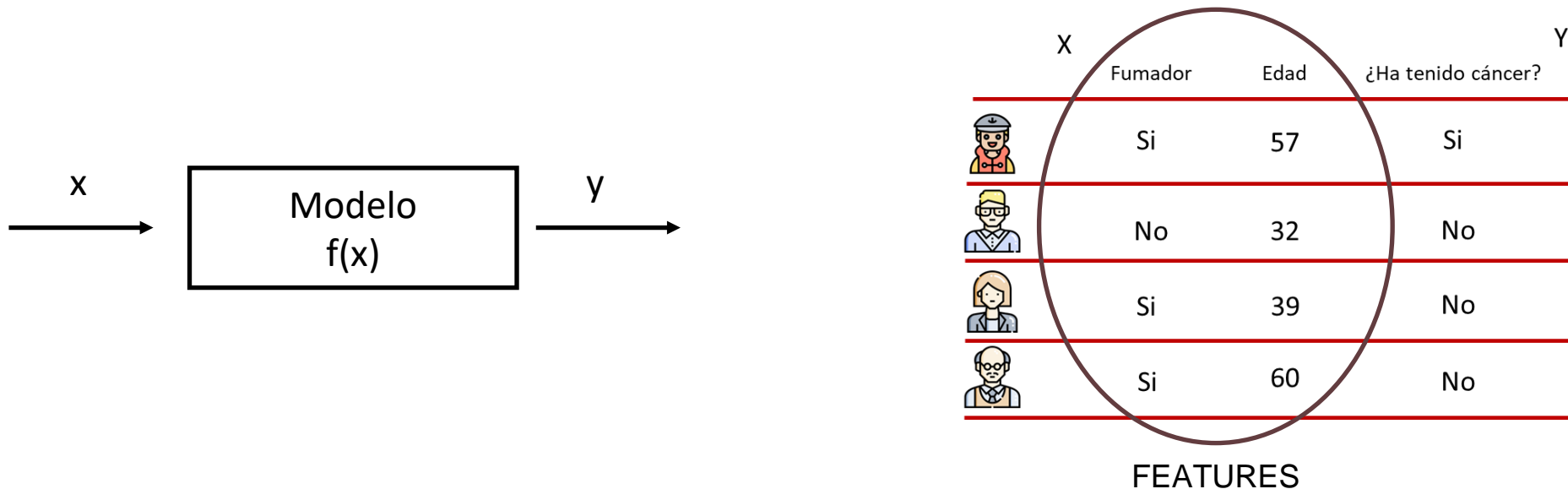
Modelos: Target (y) y Features (x)

Target (Y)

Se trata de la variable objetivo, la variable a predecir, el output del modelo. Por norma general es una única variable, ya sea numérica o categórica

Features (X)

El resto de las variables de un dataset utilizadas como input del modelo (para aprender y para predecir)



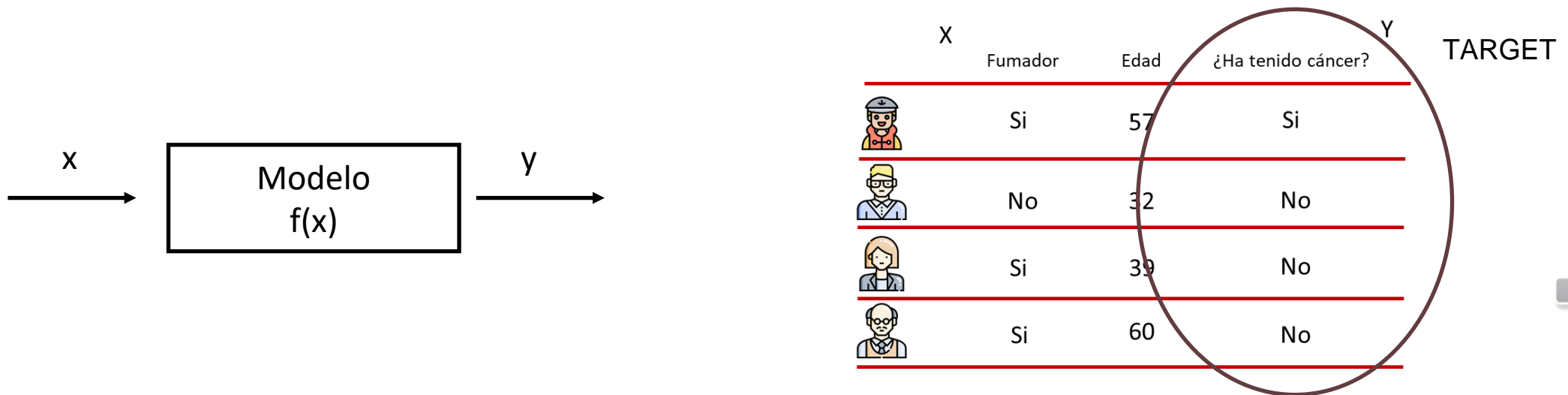
Modelos: Target (y) y Features (x)

Target (Y)

Se trata de la variable objetivo, la variable a predecir, el output del modelo. Por norma general es una única variable, ya sea numérica o categórica

Features (X)

El resto de las variables de un dataset utilizadas como input del modelo (para aprender y para predecir)



Modelos: Target (y) y Features (x)

Target (Y)

Se trata de la variable objetivo, la variable a predecir, el output del modelo. Por norma general es una única variable, ya sea numérica o categórica

Features (X)

El resto de las variables de un dataset utilizadas como input del modelo (para aprender y para predecir)

	X			Y
	Fumador	Edad	Sexo	¿Ha tenido cáncer?
	No	17	H	?
	Si	41	H	?
	Si	26	M	?
	Si	60	H	?

(Será el target aunque seguramente cambiaríamos el nombre por algo como “Alta probabilidad de Cáncer”)



Modelos: Target (y) y Features (x)





Target (Y)

Se trata de la variable objetivo, la variable a predecir, el output del modelo. Por norma general es una única variable, ya sea numérica o categórica

Features (X)

El resto de las variables de un dataset utilizadas como input del modelo (para aprender y para predecir)

Las features son
“Fumador” y “Edad”
 (“Sexo” no se ha
utilizado para entrenar
y el modelo no la
usará como variable)

	X			Y
	Fumador	Edad	Sexo	¿Ha tenido cáncer?
	No	17	H	?
	Si	41	H	?
	Si	26	M	?
	Si	60	H	?

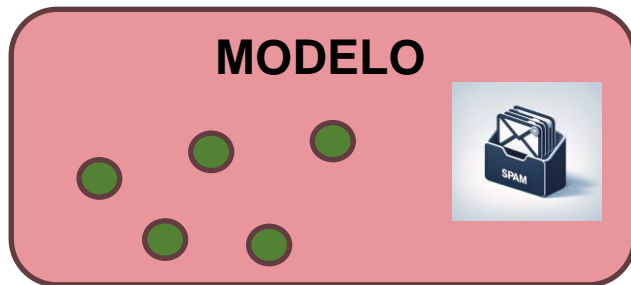


Modelos: Parâmetros e hiperparâmetros



Modelos: Parámetros e hiperparámetros

PARAMETROS

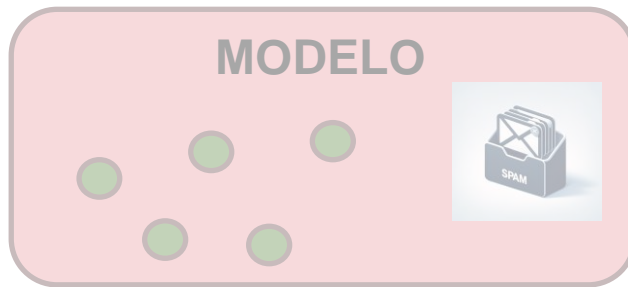


Un **parámetro** en un modelo de Machine Learning es una variable interna cuyo valor es aprendido a partir de los datos durante el entrenamiento del modelo. Estos parámetros son la esencia del modelo, ya que definen la relación entre los datos de entrada y la predicción.



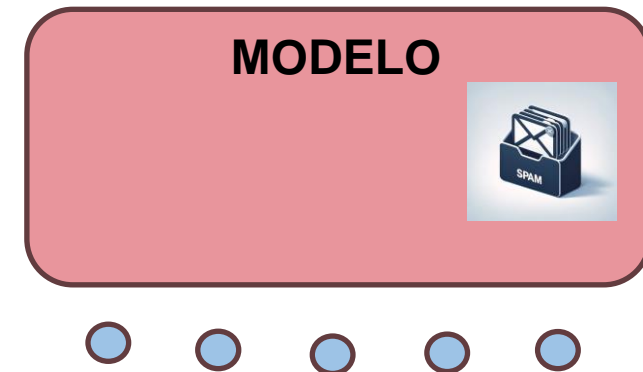
Modelos: Parámetros e hiperparámetros

PARAMETROS



Un **parámetro** en un modelo de Machine Learning es una variable interna cuyo valor es aprendido a partir de los datos durante el entrenamiento del modelo. Estos parámetros son la esencia del modelo, ya que definen la relación entre los datos de entrada y la predicción.

HIPERPARAMETROS

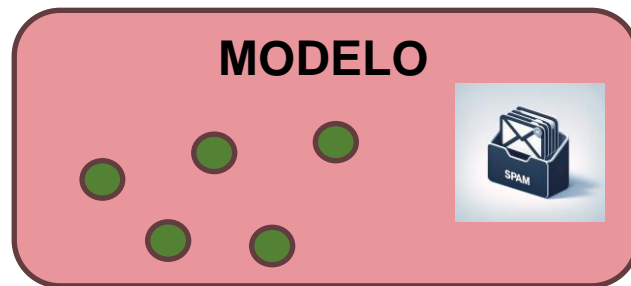


Un **hiperparámetro**, en cambio, es una configuración externa del modelo que no se aprende a partir de los datos, sino que se establece antes del proceso de entrenamiento.



Modelos: Parâmetros e hiperparâmetros

PARAMETROS



HIPERPARAMETROS

