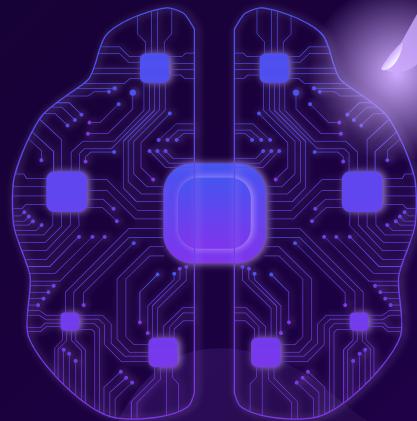


De Datos a Decisiones: Explorando IA y Machine Learning





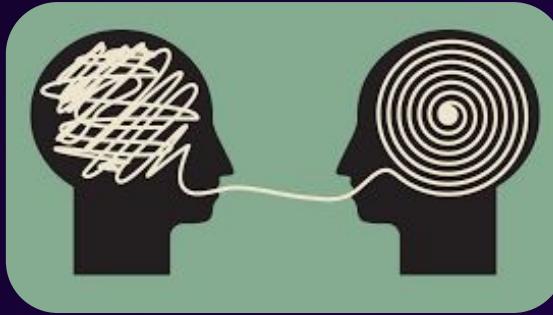
¿Qué es
una IA?

¿Qué es una Inteligencia Artificial (IA)?

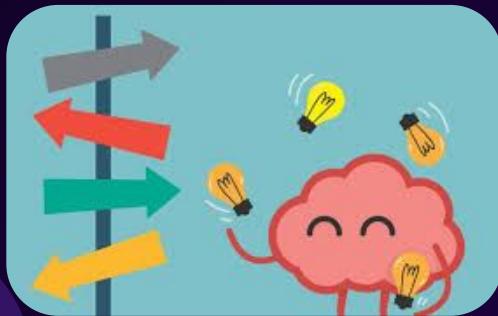
Rama de la informática que busca crear sistemas o máquinas capaces de realizar tareas que normalmente requieren **inteligencia humana**, pero no necesariamente piensa o razona como una persona



Reconocimiento de voz



Razonamiento



Toma de decisiones



Procesamiento del lenguaje natural



¿De qué se alimenta
una Inteligencia
Artificial?



Una IA se alimenta de datos

Él es **IAn**





¿Machine Learning?

mentimeter.com

5176 186



¿Qué es Machine Learning (ML)?

Aprendizaje automático, enfoque dentro de la IA centrado en el desarrollo de **algoritmos** que permiten a las máquinas **aprender** a partir de **datos**

Analizan patrones, toman decisiones y predicen sin intervención humana directa



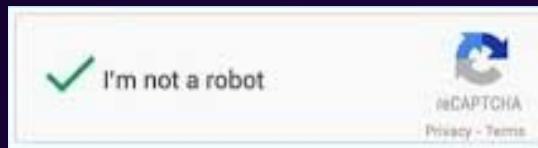
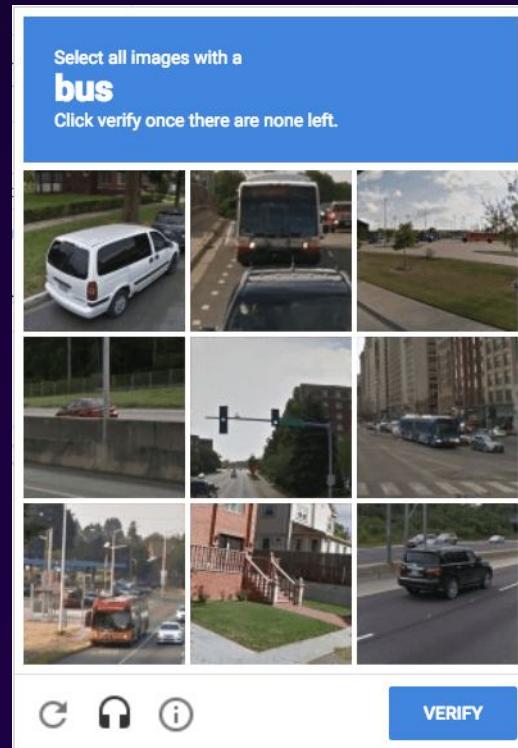
¿Cómo aprenden las máquinas?



Aprendizaje supervisado

Se entrena con un conjunto de **datos**, donde para cada entrada hay una salida correcta asociada

Aprende la **relación** entre estos datos para **predecir** nuevas entradas en el futuro



Aprendizaje NO supervisado



Busca **patrones** o **estructuras ocultas** dentro de los datos

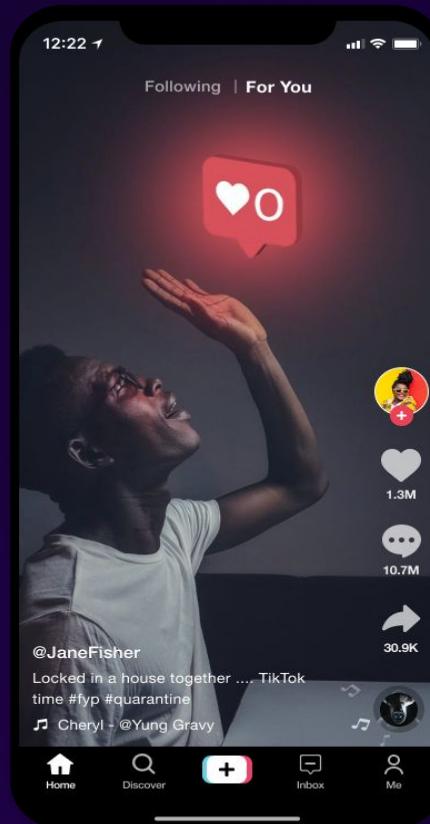
Solo conocemos los de entrada
y la **salida** es **interpretable**



Spotify®

Aprendizaje por refuerzo

Aprende a través de **prueba y error** en un entorno, recibiendo **premios** o **castigos** por sus acciones

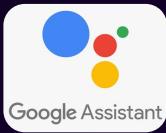




IA en nuestro
día a día

Asistentes virtuales

01



Sistemas de recomendación

03



Arte

02



Reconocimiento facial

04



Optimización del tráfico en mapas



05

Conducción autónoma



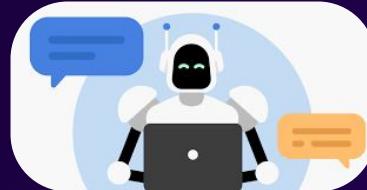
07

Traducción automática



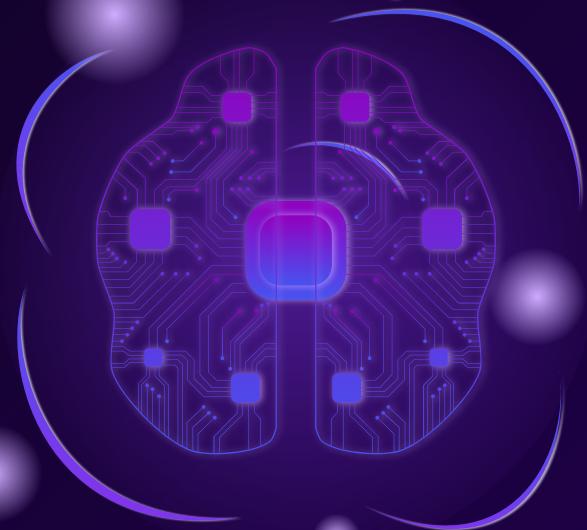
06

Chatbots



08

El camino de los datos a las decisiones



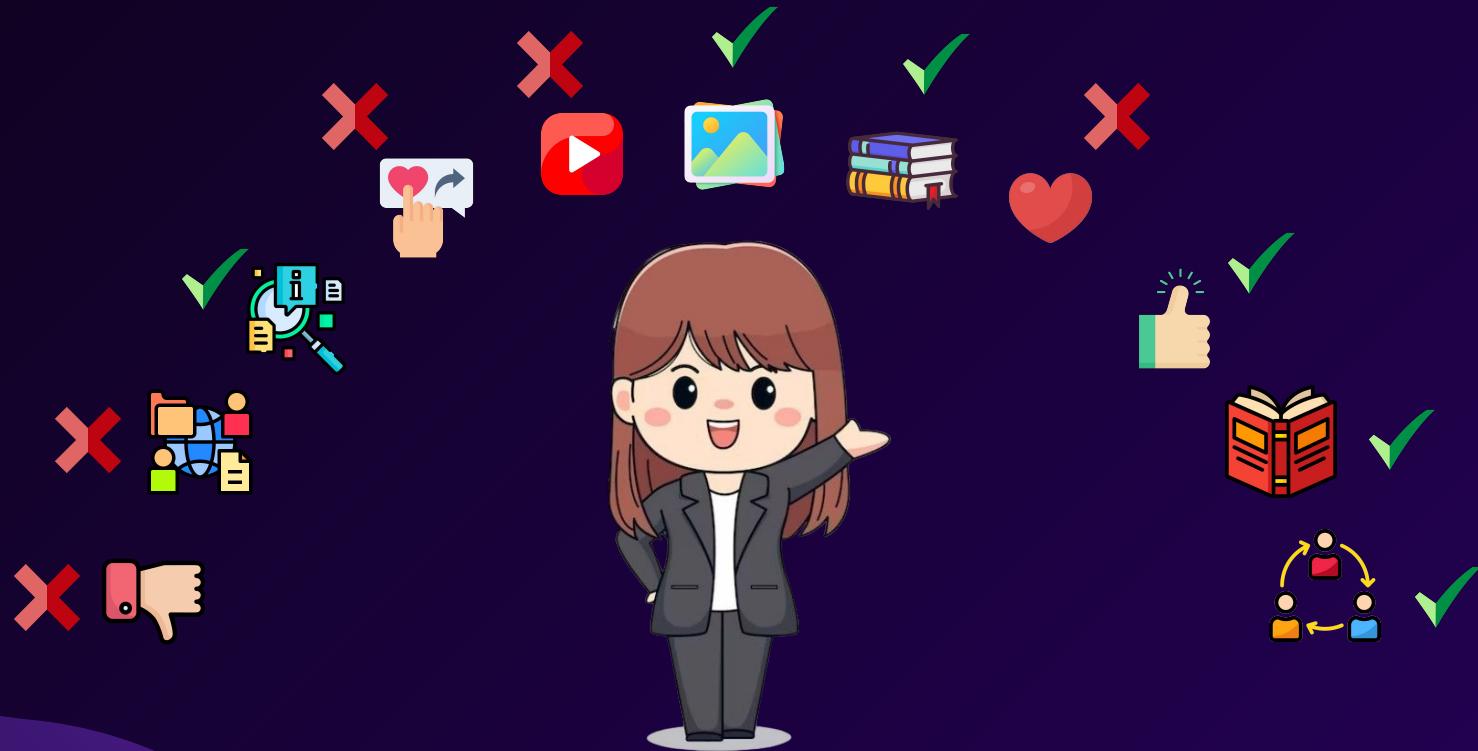


¿Se acuerdan de IAn
come datos?

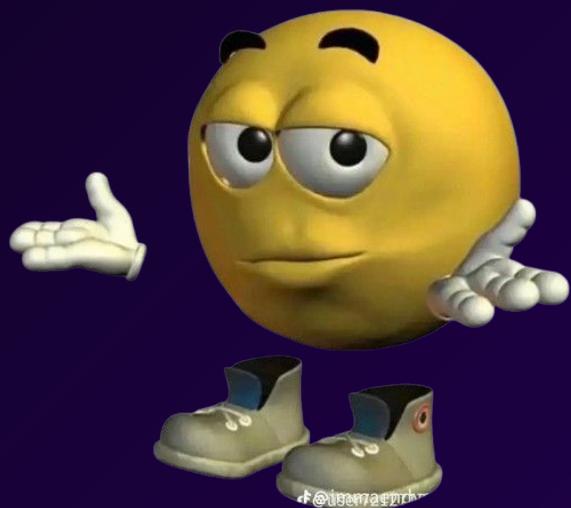
¿Qué podemos hacer para que **IAn** no se enferme?



Tenemos que **seleccionar** los datos que **IAn** va a procesar



¿Qué sigue?



@ismaraym

Vamos a limpiar
los **datos** que
seleccionamos



¿Y después?



Tenemos que
tener **claro** el
objetivo de
nuestro modelo
de IA



Algoritmos de aprendizaje

Árboles de decisión



Redes neuronales convolucionales



K-Means (K-Medias)



Gradient Boosting



Regresión lineal y logística



Aprendizaje profundo



Clustering



Q-Learning



The background features a dark purple gradient with two semi-transparent hands, one on each side, pointing towards the center. Numerous small, glowing purple dots of varying sizes are scattered across the background.

Veamos un ejemplo



1. *Divide los datos*
2. *Hace preguntas*
3. *Predice un número*

Un **árbol de decisión** es:

Modelo de **predicción** utilizado en el campo IA y ML, toma **decisiones** en función de un conjunto de **condiciones** establecidas, representadas **gráficamente** de manera **jerárquica** similar a un árbol

tinyurl.com/EjemploArbolDecision

MODELO DE PREDICCIÓN DE RIESGO DE CÁNCER DE CUELLO UTERINO

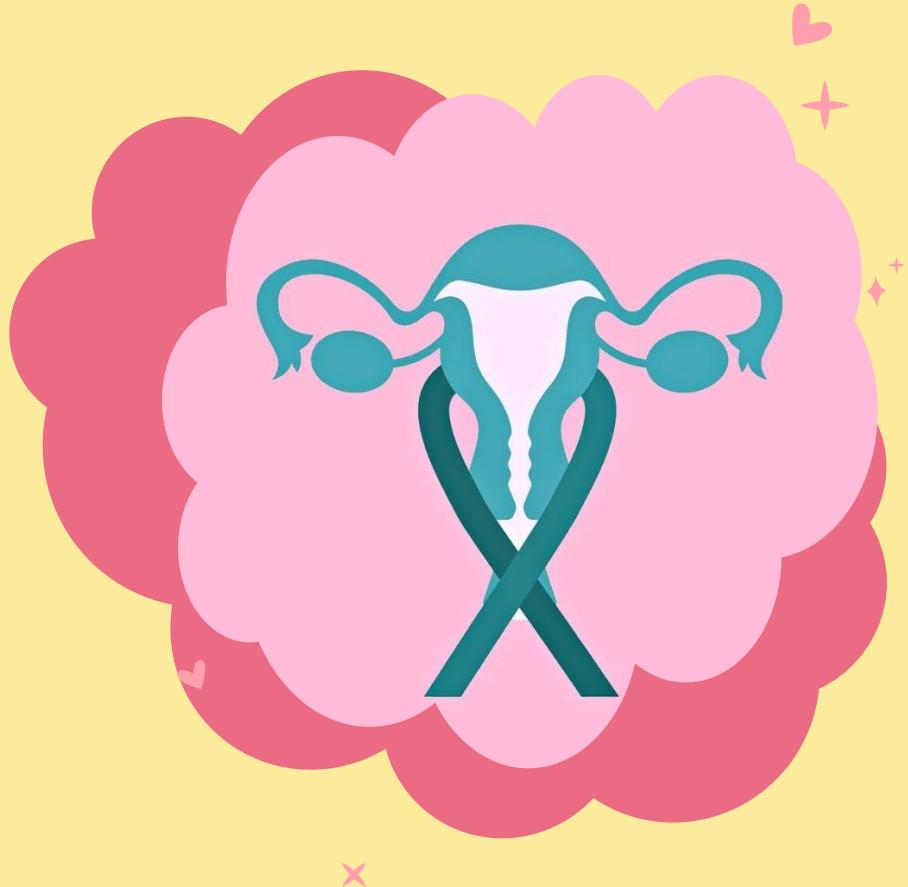


¿QUÉ ES EL CÁNCER DE CUELLO UTERINO?

El cuello uterino es la parte más baja y estrecha que sirve de entrada al útero, es aquí donde las células comienzan a multiplicarse sin control

Entre sus factores de riesgo están:

- Actividad sexual a temprana edad
- Varias parejas sexuales
- Diferentes ITS incluyendo al VPH
- Tabaquismo





¿Estamos conscientes de la
posibilidad de **contraerlo**?

¿Cuáles son sus **probabilidades**?



660,00

Mujeres que fueron
diagnosticadas con cáncer de
cuello uterino en 2022

47,199

Casos nuevos por
año

350,000

Muertes por cáncer
cervical en el año 2022



PROBLEMAS

A RESOLVER



Cómo podemos saber por diferentes antecedentes referentes a malos diagnósticos, el cáncer cérvico uterino comienza con una infección viral que normalmente aparece antes de los 30 años, al no ser tratada crea una displasia que da lugar a un **pre cáncer** que suele matar a sus portadoras en ***menos de 5 años.***



RESOLUCIÓN Y HERRAMIENTAS



OBTENCIÓN DEL DATASET

¿Dónde obtuvimos el dataset?

Kaggle

Cervical
Cancer

Edad, ITS's,
Anticonceptivos
hormonales, etc.

36
columnas

858
filas

¿Por qué éste dataset?

Datos numéricos
conforme a las
columnas y cada
paciente

HERRAMIENTAS UTILIZADAS

colab



NumPy

pandas



seaborn

VARIABLES IMPLEMENTADAS

- EDAD
- NRO. DE PAREJAS SEXUALES
- PRIMERA RELACIÓN SEXUAL
- NRO. DE EMBARAZOS



- ANTICONCEPTIVOS HORMONALES
- USO DE DIU (DISPOSITIVO INTRAUTERINO)
- INFECCIONES DE TRANSMISIÓN SEXUAL (ITS'S)
- DIAGNÓSTICO

MÉTRICAS DE PRECISIÓN DE LAS PREDICCIONES

PASO 1

PASO 2

PASO 3

PASO 4

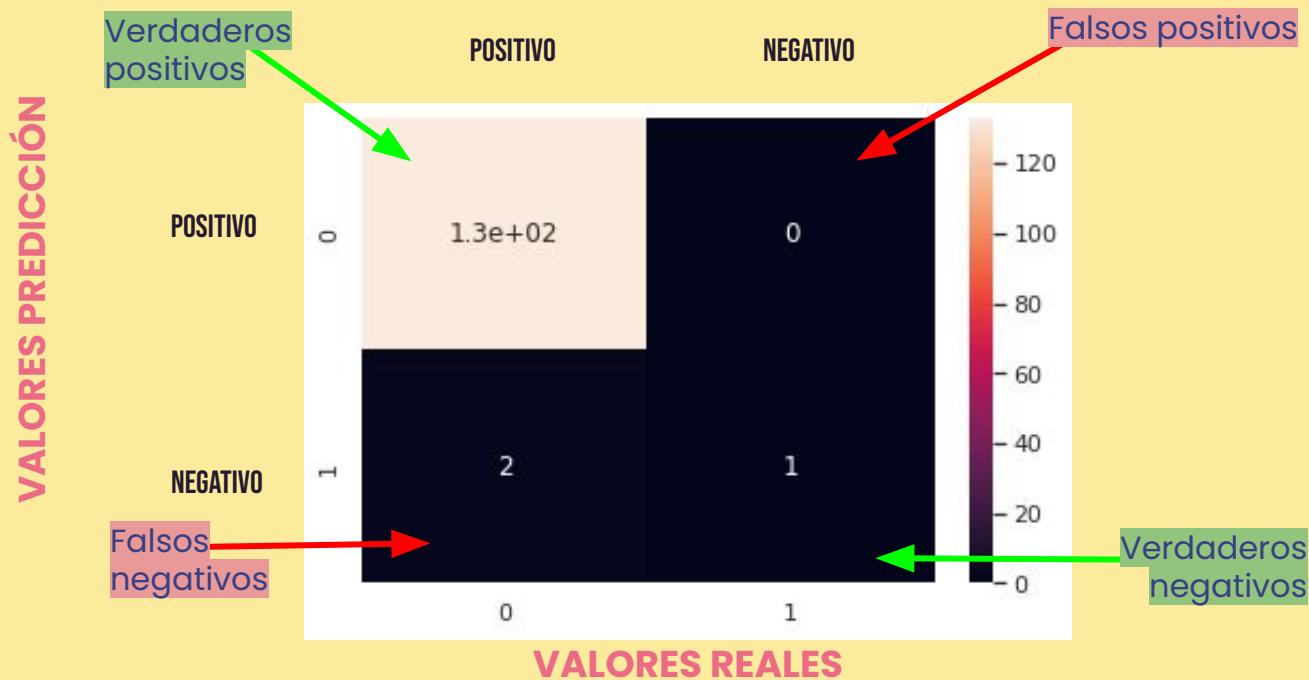
PASO 5

PASO 6

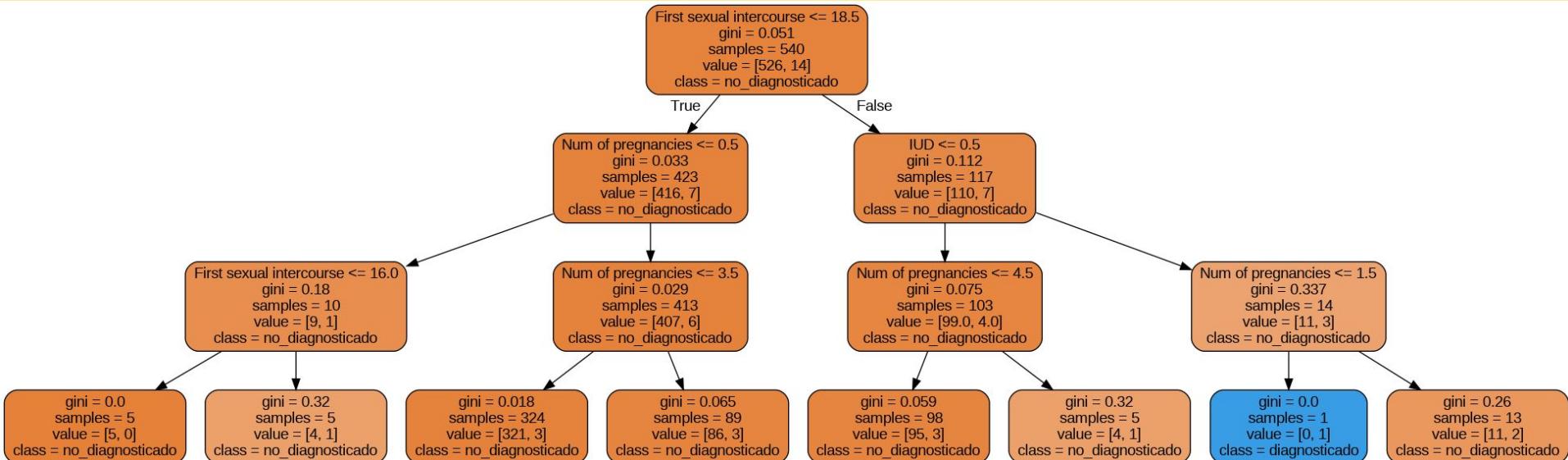
PASO 7

Definir las variables	Separar los datos	Modelo	Entrenar	Predictión	Precisión	Matriz de confusión y árbol de decisión
Definimos las variables a implementar	Separa los datos dados con los datos objetivo	Niveles que se mostrarán en el árbol de decisión	Entrena al algoritmo mediante los datos dados	Predice la probabilidad de que ocurra algo	Mide la eficiencia del algoritmo	Permite visualizar el desempeño del algoritmo

MATRIZ DE CONFUSIÓN



ÁRBOL DE DECISIÓN

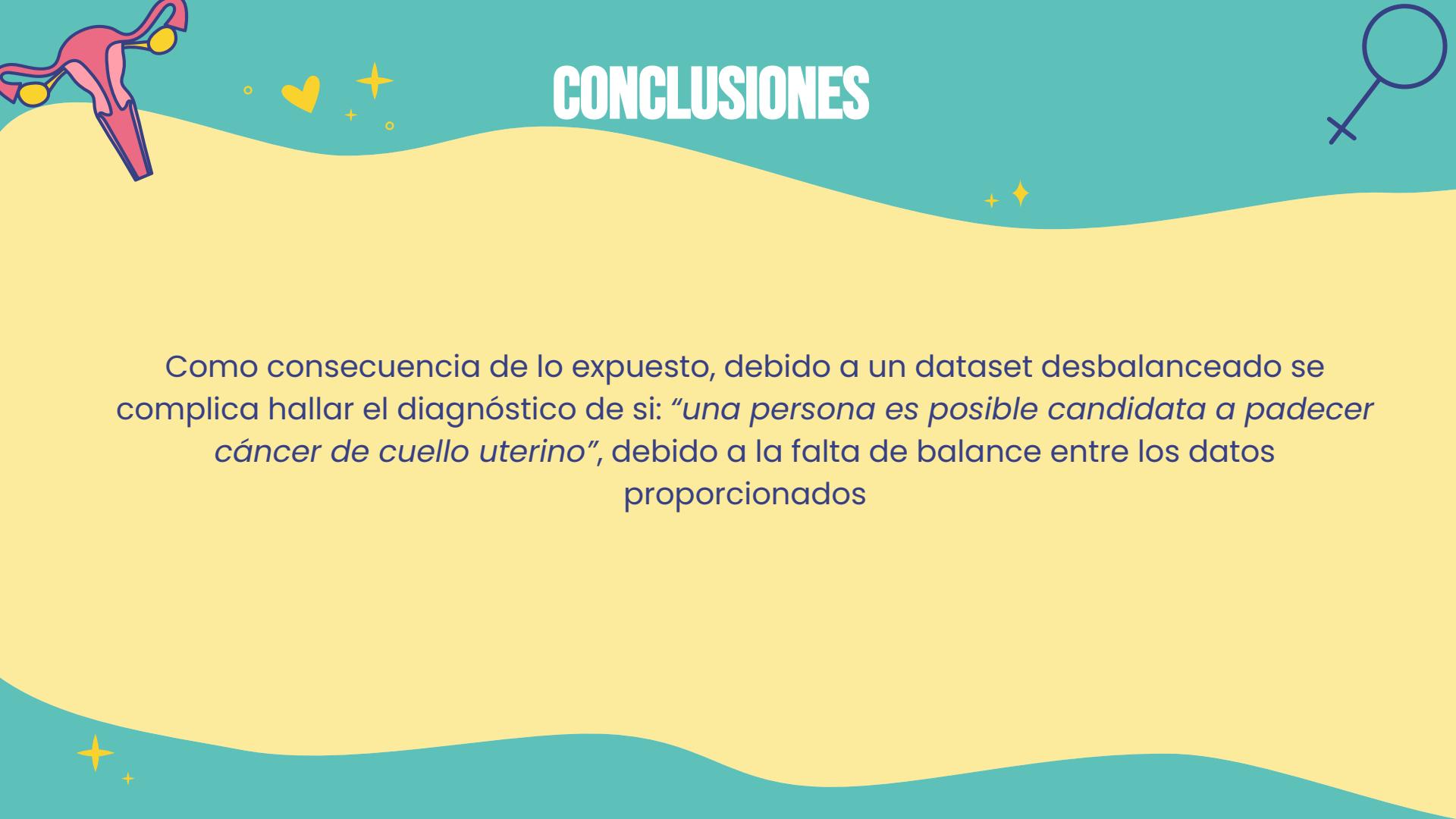


MÉTRICAS DE PRECISIÓN DE LAS PREDICCIONES

0.9852941176470

De forma que el modelo planteado con las variables seleccionadas, resulta con una precisión favorable





CONCLUSIONES

Como consecuencia de lo expuesto, debido a un dataset desbalanceado se complica hallar el diagnóstico de si: *"una persona es posible candidata a padecer cáncer de cuello uterino"*, debido a la falta de balance entre los datos proporcionados

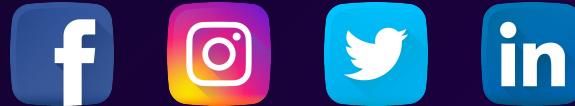
¿Qué nos diferencia de
una IA?

CURIOSIDAD

¡Siempre sean
curiosos(as) !

¡Muchas gracias!

*Que tengan un **lindo** día :)*



@dana_asturizaga