



**INTELIGENCIA ARTIFICIAL  
PROYECTO ENTREGA #1**

**DOCENTE  
RAUL RAMOS POLLAN**

**ESTUDIANTES  
ADRIANA ISABEL RIOS  
MARIA CAMILA LOPERA  
SARA PÉREZ HIGUITA**

**HORARIO MARTES-JUEVES 10-12 PM.**

**UNIVERSIDAD DE ANTIOQUIA  
JULIO DEL 2022**

## **1. Problema predictivo por resolver.**

El cuerpo humano tiene mecanismos de defensa que generan diferentes respuestas en contacto con las distintas afecciones que puedan presentarse.

La sepsis es una respuesta extrema del cuerpo ante una infección de alto nivel. Es considerada una emergencia médica ya que, sin un tratamiento oportuno, puede provocar rápidamente daños en los tejidos, insuficiencia orgánica e incluso la muerte.

La detección temprana de la sepsis es esencial para mejorar el pronóstico, tiene el potencial de salvar vidas y limitar los recursos hospitalarios requeridos para atender la emergencia.

Por esto, con los valores de la sintomatología de las sepsis expresadas en el dataset a continuación, se tendrá como propósito desarrollar un modelo predictivo para predecir la sepsis tempranamente (6 horas antes de la predicción clínica) y falsas alarmas en dado caso.

## **2. Dataset a utilizar.**

El dataset fue seleccionado de la plataforma Kaggle

[<https://www.kaggle.com/datasets/salikhussaini49/prediction-of-sepsis>], que tiene 5100 Muestras de pacientes y 39 columnas que se presentan a continuación.

### **Signos vitales (columnas 1-8)**

FC Frecuencia cardíaca (latidos por minuto)

O2Sat Pulsioximetría (%)

Temp Temperatura (Grados C)

PAS PA sistólica (mm Hg)

PAM Presión arterial media (mm Hg)

PAD PA diastólica (mm Hg)

Resp Tasa de respiración (respiraciones por minuto)

EtCO2 Dióxido de carbono corriente final (mm Hg)

### **Valores de laboratorio (columnas 9-34)**

BaseExcess Medida de exceso de bicarbonato (mmol/L)

HCO3 Bicarbonato (mmol/L)

FiO2 Fracción de oxígeno inspirado (%)

pH N/A

PaCO2 Presión parcial de dióxido de carbono de la sangre arterial (mm Hg)

SaO2 Saturación de oxígeno de la sangre arterial (%)

AST Aspartato transaminasa (UI/L)

BUN Nitrógeno ureico en sangre (mg/dL)

Fosfatasa alcalina Fosfatasa alcalina (UI/L)

Calcio (mg/dL)

Cloruro (mmol/L)

Creatinina (mg/ dL)  
Bilirrubina directa Bilirrubina directa (mg/dL)  
Glucosa Glucosa sérica (mg/dL)  
Lactato Ácido láctico (mg/dL)  
Magnesio (mmol/dL) Fosfato (mg/dL)  
Potasio (mmol/L)  
Bilirrubinatotal Bilirrubina total (mg/dL)  
Troponina I Troponina I (ng/mL)  
Hct Hematocrito (%)  
Hgb Hemoglobina (g/dL)  
PTT Tiempo de tromboplastina parcial (segundos)  
Recuento de leucocitos WBC (recuento  $10^3/\mu\text{L}$ )  
Fibrinógeno (mg/ dL)  
Plaquetas (recuento  $10^3/\mu\text{L}$ )

#### **Datos demográficos (columnas 35-40)**

Edad Años (100 para pacientes de 90 años o más)  
Sexo Mujer (0) o Hombre (1)  
HospAdmTime Horas entre ingresos hospitalarios y ingreso en UCI.  
Duración de la estadía en UCI (horas desde el ingreso en UCI)

#### **Resultado (columna 39)**

SepsisLabel Para pacientes con sepsis, SepsisLabel es 1 si  $t \geq t_{\text{sepsis}}$ –6 y 0 si t

### **3. Métricas de desempeño requeridas.**

El rendimiento del modelo se evaluará utilizando datos de discriminación y calibración independientes.

Para la evaluación del desempeño del modelo se utilizará el puntaje F1, que incluye las métricas de precisión, y sensibilidad; lo cual resulta conveniente para el contexto clínico trabajado, ya que en este caso la distribución de las clases es desigual.

### **4. Crítica sobre desempeño deseable en producción.**

Se utilizará una función de utilidad. El modelo de predicción de sepsis en pacientes debería de tener un porcentaje de acierto  $>90\%$ , y también un falso positivo  $<10\%$ , ya que es una patología grave en la que no conviene una predicción tardía que agravará la condición clínica ni un falso pronóstico que conllevará al desaprovechamiento de recursos hospitalarios.