# 

# **ETL - Detección de sepsis:**

1) Pacientes mayores a 14 años. (edad registrada en el momento de ingreso)

2) Pacientes con caracteristicas con mas del 60% de datos faltantes, eliminarlos. (porque si no tengo caracteristicas que necesito evaluar, esos que tienen más del 60% faltantes no me sirven)

3) Solo pacientes que tuvieran al menos 3 registros de signos vitales.

4) Excluir a los que ya ingresaron con sepsis (porque es lo que yo quiero determinar y si ya tiene sepsis no tengo que determinar nada)

# **ETL - Predicción de mortalidad después de detectar sepsis**

(es decir tener todos esos mismos datos y además):

5) Estancia en UCI mayor a 2 días. (porque si fue menor a 2 días significa que se le dio el alta y salió de complicación)

6) Más de 3 mediciones de glucosa en cada uno de los tramos 0-24hs y 24-48hs. (porque depende qué tan rápido se hayan elevado los valores, habrá mayor o menor probabilidad de muerte)

# Definición sepsis

Se define la sepsis como “una disfunción orgánica potencialmente mortal causada por una respuesta desregulada del huésped a la infección” y se propone, la inclusión de criterios para tener en cuenta en el diagnóstico de la sepsis.

# Criterios

## qSOFA: evaluar la posibilidad de riesgo para pacientes con sospecha de sepsis (0-3).

## 2 o + => Riesgo más alto de mortalidad o estancia prolongada

## Menor a 2 => No reduce la probabilidad de padecer sepsis, hay que seguir monitoreando.

-Frecuencia respiratoria >= 22 resp/min

-Alteración del sensorio < 15

-Presión arterial sistólica <= 100 mmHg

## SOFA: definir si existe disfunción o fallo orgánicos. (0-4)

## Diferente y < 3 => Hay disfunción orgánica.

## >3 => Fallo orgánico

-Respiración Pa03/FiO3 (mmHg)

-Coagulación (Plaquetas)

-Hígado (Bilirrubina mg/dl)

-Cardiovascular (PAM o su manejo)

-Sistema Nervioso Central (Escala de coma de Glasgow)

Renal (Creatina ml/dL Y Gasto urinario mL/dia)

## SAPS-II:estimacion de la gravedad y mortalidad, este evalúa los “peores valores” registrados del enfermo durante las primeras 24 horas tras su ingreso

-HR

-SBP

-TEMP

-PAO2/FIO2

-UO (ML)

-S. UREA

-TLC

-K

-Na

-BICARBONATO

-BIL

-GCS

# Datos generales a tomar:

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

# Puntuación para cada criterio

# Qsofa:

El puntaje para cada uno de los criterios de este indicador

oscila entre 0 y 3 puntos; en caso de obtener una puntuación de 2 o más puntos, se asocia a un riesgo más alto de mortalidad o estancia prolongada en UCI. Una puntuación baja, no reduce la probabilidad de que un paciente padezca sepsis, se recomienda seguir

monitoreando y evaluando el estado de este. Una puntuación alta por su parte, permite la adopción de medidas específicas de tratamiento que incluye la evaluación del puntaje

SOFA (W. Seymour et al., 2016).

# SOFA:

El punto base de SOFA es el cero, es decir, que el puntaje mínimo para los criterios que este indicador incluye es de 0 y el máximo de 4. Una puntuación diferente de 0 y menor de 3, indica la existencia de disfunción orgánica y puntajes mayores, indican fallo orgánico. El incremento en esta puntuación durante las primeras horas de ingreso, refleja un riesgo en la mortalidad superior al 49%, y una puntuación mayor de 15 puntos refleja una mortalidad mayor al 90% (Moreno et al., 1999).

En la siguiente tabla se encuentra las variables o criterios que incluye este indicador,

además los puntos asignados a cada uno:

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

SAPS-II

Puntaje, oscila entre 0 y 163 puntos. Dichos valores, son recalculados sólo si el paciente reingresa luego del alta. Este indicador se compone de 12 variables fisiológicas y 3 variables relacionadas con la enfermedad. A estas variables, se añaden otros criterios la salud previa del paciente, y datos demográficos como la edad. Con todo ello, y dependiendo de la desviación de cada una de las variables con respecto de los valores normales, se analizan mediante una ecuación logística, que permite evaluar desde el punto de vista estadístico, la probabilidad de muerte del paciente (Le Gall, Lemeshow, & Saulnier, 1993).

Tabla

Descripción generada automáticamente

* Se realizó la búsqueda de las variables necesarias para ejecutar la medición de los modelos de severidad SOFA, qSOFA y SAPSII, estos modelos son utilizados para poder estimar sepsis, tiempo de estancia en la UCI y mortalidad.
* Luego de localizar los datos se utilizaron varios procesos ETL para unir los datos y así logra que cada entrada al hospital (Hadm\_id) tuviera asociada las variables de los modelos. La unión se realizó por medio de 3 ETL, para cada modelo, utilizando la clave doble: Subject\_id y Hadm\_id.
* Adicionalmente, al unir por ambas claves se logra que cada estancia en la UCI, sin importar si un paciente ingresa más de una vez, tenga todas las variables de los modelos.

DATOS DE PREDICCIÓN (O CLASIFICACIÓN):

* Sepsis explícita
* Predicción de Angus
* Estancia en pacientes
* Muerte (Si la fecha de fallecimiento del paciente es igual o menor a la fecha de salida de la UCI, entonces el paciente falleció en ella.
* Exportar a Excel, armar una matriz de confusión de sepsis:
* Esto es para obtener la eficiencia de los modelos, donde se calcula la matriz de confusión de los métodos para obtener la precisión de las predicciones de sepsis explicita, Angus y mortalidad.
* Matriz de confusión de mortalidad.
* Teniendo todo unificado, separo en datos a entrenar y datos para validar.
* Aplico los modelos:

# Modelo de Machine Learning a seguir:

DE CLASIFICACIÓN

* KNN K-Nearest Neighbors
* SVM SV Machine
* Naive Bayes
* Decision Tree
* Random Forest

# Evaluación o prueba del modelo:

* Validación cruzada
* Cálculo de accuracy, recall, precisión. Y se ajustan los modelos usando validación cruzada:

Interfaz de usuario gráfica, Texto, Aplicación, Carta

Descripción generada automáticamente

Se evaluó: precisión del clasificador y la pérdida de registro del clasificador.

Preveer si el paciente fallecerá en la UCI.