

Universidad Católica "Nuestra señora de la Asunción"  
Facultad de ciencias y tecnologia  
Complementos de Informatica



## Propuesta de Problemática y Solución

Entrega 1 - Data mining

**Alumno:** Alain Vega

**Profesor:** Wilfrido Felix Inchausti Martínez

11 de septiembre de 2023

# Índice

<b>1. Caso de estudio: Insuficiencia cardiaca</b>	<b>2</b>
1.1. Problema . . . . .	2
<b>2. Determinar objetivos del negocio</b>	<b>2</b>
2.1. Background . . . . .	2
2.2. Objetivos del negocio . . . . .	2
2.3. Criterios de exito del negocio . . . . .	3
<b>3. Valoracion de la situacion</b>	<b>3</b>
3.1. Inventario de recursos . . . . .	3
3.2. Requisitos, supuestos y restricciones . . . . .	3
3.3. Riesgos y Contingencias . . . . .	3
3.4. Costes y Beneficios . . . . .	4
<b>4. Determinar los objetivos del DM</b>	<b>4</b>
4.1. Metas de Data Mining . . . . .	4
4.2. Criterios de éxito de DM . . . . .	4
<b>5. Bibliografia</b>	<b>5</b>

# 1. Caso de estudio: Insuficiencia cardíaca

## 1.1. Problema

Las enfermedades cardiovasculares (ECV) son la principal causa de muerte a nivel mundial, cobrando la vida de aproximadamente 17.9 millones de personas cada año, lo que representa el 31 % de todas las muertes en todo el mundo. Cuatro de cada 5 muertes por ECV son causadas por ataques cardíacos y accidentes cerebrovasculares, y un tercio de estas muertes ocurren prematuramente en personas menores de 70 años.

La insuficiencia cardíaca es un evento común causado por las ECV. Las personas con enfermedades cardiovasculares o que tienen un alto riesgo cardiovascular (debido a la presencia de uno o más factores de riesgo como hipertensión, diabetes, hiperlipidemia o enfermedades ya establecidas) necesitan una detección temprana y gestión, en la cual un modelo de aprendizaje automático puede ser de gran ayuda.

# 2. Determinar objetivos del negocio

## 2.1. Background

La insuficiencia cardíaca se produce cuando el músculo cardíaco no bombea sangre de la manera que debería. Cuando esto sucede, la sangre a menudo retrocede y el líquido se puede acumular en los pulmones, lo que causa falta de aliento.

El conjunto de datos a considerar, analiza 11 características de un total de 918 personas.

## 2.2. Objetivos del negocio

- **Reducción de costos de atención médica:** al prevenir diagnósticos fallidos y tratamientos innecesarios, se reducen los costos asociados con el tratamiento de insuficiencia cardíaca y enfermedades relacionadas.
- **Mejorar la atención médica:** al identificar de manera más efectiva a aquellos en riesgo de insuficiencia cardíaca. Esto conlleva a un tratamiento más oportuno y eficaz.

- **Aumentar la calidad de vida de los pacientes:** al permitirles recibir atención preventiva y seguimiento temprano gracias a la predicción.

### 2.3. Criterios de éxito del negocio

- **Reducción de costos:** disminuir los costos de tratamientos médicos un 20 % para los pacientes
- **Impacto en la vida de los pacientes:** afectar positivamente la calidad de vida de los pacientes.

## 3. Valoración de la situación

### 3.1. Inventario de recursos

Se cuenta con los siguientes datos médicos: *edad, sexo, tipo de dolor en el pecho, presión arterial en reposo, colesterol, azúcar en sangre en ayunas, resultados del electrocardiograma en reposo, frecuencia cardíaca máxima alcanzada, angina inducida por el ejercicio, pico antiguo = ST, pendiente del segmento ST durante un ejercicio físico máximo en una prueba de esfuerzo cardíaco, sufrió de insuficiencia cardíaca.*

Que sirvan para el entrenamiento y pruebas del modelo.

### 3.2. Requisitos, supuestos y restricciones

- **Requisitos:** detección temprana de posible insuficiencia cardíaca con una tasa de falsos negativos cada vez menor en proporción a la cantidad de pacientes analizados.
- **Supuestos:** los datos iniciales proveídos son de fiar, representan las características principales a analizar.
- **Restricciones:** la fiabilidad del modelo nunca debe bajar del 90 %

### 3.3. Riesgos y Contingencias

- **Riesgos:** algunas afecciones cardíacas progresivamente dejan el corazón demasiado débil o rígido como para llenarse y bombear sangre

de forma apropiada. Por lo que los síntomas pueden desarrollarse lentamente. Sin embargo a veces, los síntomas de la insuficiencia cardíaca comienzan repentinamente.

Entonces existe el riesgo de que el conjunto de datos a analizar no contemple todas las variables para un análisis exacto.

- **Contingencias:** agregar nuevas características relevantes al conjunto de datos, además de incluir más registros con el fin de tener una muestra más grande y variada.

### 3.4. Costes y Beneficios

De acuerdo a un estudio realizado en Estados Unidos, el costo promedio por año de pacientes con insuficiencia cardíaca en dicho país es de **\$18.953** de los cuales el 42,8 % (\$8114) son costos hospitalarios.

Gracias a una detección temprana es posible evitar que el paciente se hospitalice por un cuadro avanzado de esta afección. Por lo que es posible reducir el costo anual a  $\$18.953 * (1 - 0,428) = \$10.841,116$ .

Esto es una reducción del 57,2 % de los costos promedios anuales en el tratamiento.

## 4. Determinar los objetivos del DM

### 4.1. Metas de Data Mining

- **Precision del modelo:** obtener un modelo fiable, que posea una precisión mayor o igual al 95 %
- **Implementación Clínica:** implementar el modelo en más de un entorno clínico real para ayudar en la detección temprana y la gestión de la insuficiencia cardíaca.

### 4.2. Criterios de éxito de DM

Se establecen como criterios de éxito, los siguientes:

- **Precision del modelo:** obtener un modelo fiable, que posea una precisión de por lo menos el 90 %

- **Reduccion de costos medicos:** disminuir los costos de tratamientos medicos un 57,2 % para los pacientes

## 5. Bibliografia

- Metodologia CRISP-DM
- Conjunto de datos de insuficiencia cardiaca - *kaggle*
- Costo promedio anual de la insuficiencia cardiaca