



## PROYECTO “Patient Survival Prediction”

Miguel Angel Urueña Riobo

[miguel.uruena@udea.edu.co](mailto:miguel.uruena@udea.edu.co)

C.C 1.006.121.797

Gaia Ramirez Hincapíe

[gaia.ramirez@udea.edu.co](mailto:gaia.ramirez@udea.edu.co)

C.C 1.005.273.358

Tutor:

Raul Ramos Pollan

[raul.ramos@udea.edu.co](mailto:raul.ramos@udea.edu.co)

INTRODUCCIÓN A LA INTELIGENCIA ARTIFICIAL

CÓDIGO: 2508401 - INGENIERÍA DE SISTEMAS V4

**Nombre del Proyecto:** Patient Survival Prediction

### Links Github:

- [https://github.com/uruenariobo/ai4eng\\_health\\_issues](https://github.com/uruenariobo/ai4eng_health_issues)
- [https://github.com/gaiamilenium99/ai4eng\\_health\\_issues1](https://github.com/gaiamilenium99/ai4eng_health_issues1)

### 1. Descripción del problema predictivo a resolver:

Dadas las características de un paciente y su historia clínica vamos a predecir la probabilidad de supervivencia de dicho paciente.

### 2. Dataset a utilizar:

Vamos a usar el dataset de kaggle “Patient Survival Prediction” ([enlace](#)), que tiene 91.713 número de muestras (pacientes) y 85 columnas, de las cuales las columnas más representativas son:

Edad.

Índice de masa corporal.

Género.

Peso.

Presión sanguínea

Ritmo cardíaco máximo.



La ubicación del paciente antes de su ingreso en la unidad.

El APACHE IV (es una predicción probabilística de la mortalidad hospitalaria del paciente).

### **3. Métricas de desempeño requeridas (de machine learning y de negocio):**

Nuestro modelo de predicción de la supervivencia en pacientes debería de tener un porcentaje de acierto  $>80\%$ , pero también un false negative rate  $<10\%$ , ya que es un dictamen grave y delicado, por lo cual es preferible no fallar una predicción de un paciente que verdaderamente morirá, aunque eso implique que aumente el número de falsos positivos.

### **4. Criterio sobre cuál sería el desempeño deseable en producción:**

Si la ocupación hospitalaria de las unidades de cuidados intensivos no aumentan más de un  $10\%$  no merece la pena poner el modelo en operatividad ya que el coste de desarrollo y mantenimiento no cubriría las ganancias adicionales de ese aumento.