

Vamos a empezar a pensar en el tipo de proyecto que estamos trabajando.

Ahora ya los conoces y puedes trabajar con los siguientes:

- Proyecto 1 - descriptivo / dashboard
- Proyecto 2 – exploratorio/investigativo:
 - Comparativo
 - Relacional
 - Causal
- Proyecto 3 – clasificador
- Proyecto 4 – regresión
- Proyecto 5 – clustering
- Proyecto 6 – reducción dimensional
- Proyecto 7 – reglas de asociación
- Proyecto 8 – series temporales

Te propongo los siguientes datasets:

- <https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset>
- <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction>
- <https://www.kaggle.com/datasets/arashnic/hr-analytics-job-change-of-data-scientists>
- <https://www.kaggle.com/datasets/benroshan/factors-affecting-campus-placement>
- <https://www.kaggle.com/datasets/ayessa/salary-prediction-classification>

Elige alguno de ellos y ubícalo en uno o máximo dos proyectos tipo.

1. ¿En qué tipo de proyecto te estás enfrentando? Marca máximo dos tipos de proyectos.

- **Proyecto 1** – exploratorio/investigativo:
 - Comparativo
 - Relacional
 - Causal
- Proyecto 2 – clasificador
- **Proyecto 3** – regresión
- Proyecto 4 – clustering
- Proyecto 5 – reducción dimensional
- Proyecto 6 – reglas de asociación
- Proyecto 6 – series temporales

¡Excelente!

El siguiente paso es pensar en cómo va a ser la tabla de datos.

Sigue con el mismo dataset y trata de realizar el siguiente ejercicio:



2. Te planteo que piensas en **la tabla de datos**, que la dibujes y que ordenes las ideas de la siguiente manera:

Características clave de tu tabla de datos

- Responde a las siguientes preguntas para plantear correctamente la estructura de tu clasificador. <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction>

Tu variable respuesta o variables respuesta

Riesgo de ataque al corazon en funcion:

de la edad

Sexo del paciente

Nivel de colesterol

Tus variables de entrada. Features/Inputs

Edad

Sexo del paciente

Nivel de colesterol

¿Qué son las filas u observaciones?	pacientes
¿Qué grupos de variables o factores vas a incluir? Estos factores son los que nos darán información de tu respuesta.	- Variables relacionadas con las características del paciente. Trata de hacer una agrupación por conceptos, por factores. Dentro de cada grupo factor habrá diferentes variables que den esa característica. -... -...
Lista las variables de entrada que crees que pueden tener un impacto importante	-..Edad -...Colesterol -...Sexo del paciente

¡Gran trabajo!

Ahora ya conoces el tipo de proyecto que estás trabajando.

También conoces cómo están estructuradas las variables, cuál es la respuesta y las variables de entrada.

Fíjate que nos estamos enfocando realmente bien

¡Seguimos por buen camino!

Celébralo como más te guste.

Salgo a correr y hago ejercicio para celebrar esta hoja de trabajo ;)

1 abrazo y hasta pronto. Brindo por tus éxitos ;)

Jordi Ollé