Proyecto: Machine Learning

instituto tecnológico de morelia

Proyecto – UNIDAD 3.

M.I.A. WILFRIDO CORTES OROZCO

2023

|  |  |  |
| --- | --- | --- |
| Nombre del Alumno | Número de control | Número de Proyecto |
|  |  |  |

# Descripción y objetivos del proyecto.

Objetivos:

* Que el estudiante pueda emplear técnicas de ML para realizar predicciones.
* Que el estudiante siga reforzando sus conocimientos en el análisis y exploración de los datos.
* Que el estudiante siga realizando actividades que pueda subir en su Git para crear un portafolio de proyectos.

Descripción de la actividad:

Este proyecto consiste en cuatro etapas principales:

1. EDA.
2. Manipulación y tratamiento de los datos.
3. Machine Learning
4. Presentación del proyecto.

El dataset a utilizar se llama “cirrhosis.csv, en la *figura 1* se muestra la información del dataset.

Las instrucciones de este proyecto vienen en el archivo de Jupyter-Notebook, el trabajo debe ser realizado sobre ese archivo y debe ser subido a GitHub en un repositorio nuevo, el repositorio debe ser llamado “ML\_PROYECTO\_NombreEstudiante”, es importante llevar ese formato ya que, posteriormente en el curso, se utilizará para mostrar en el portafolio de proyectos.

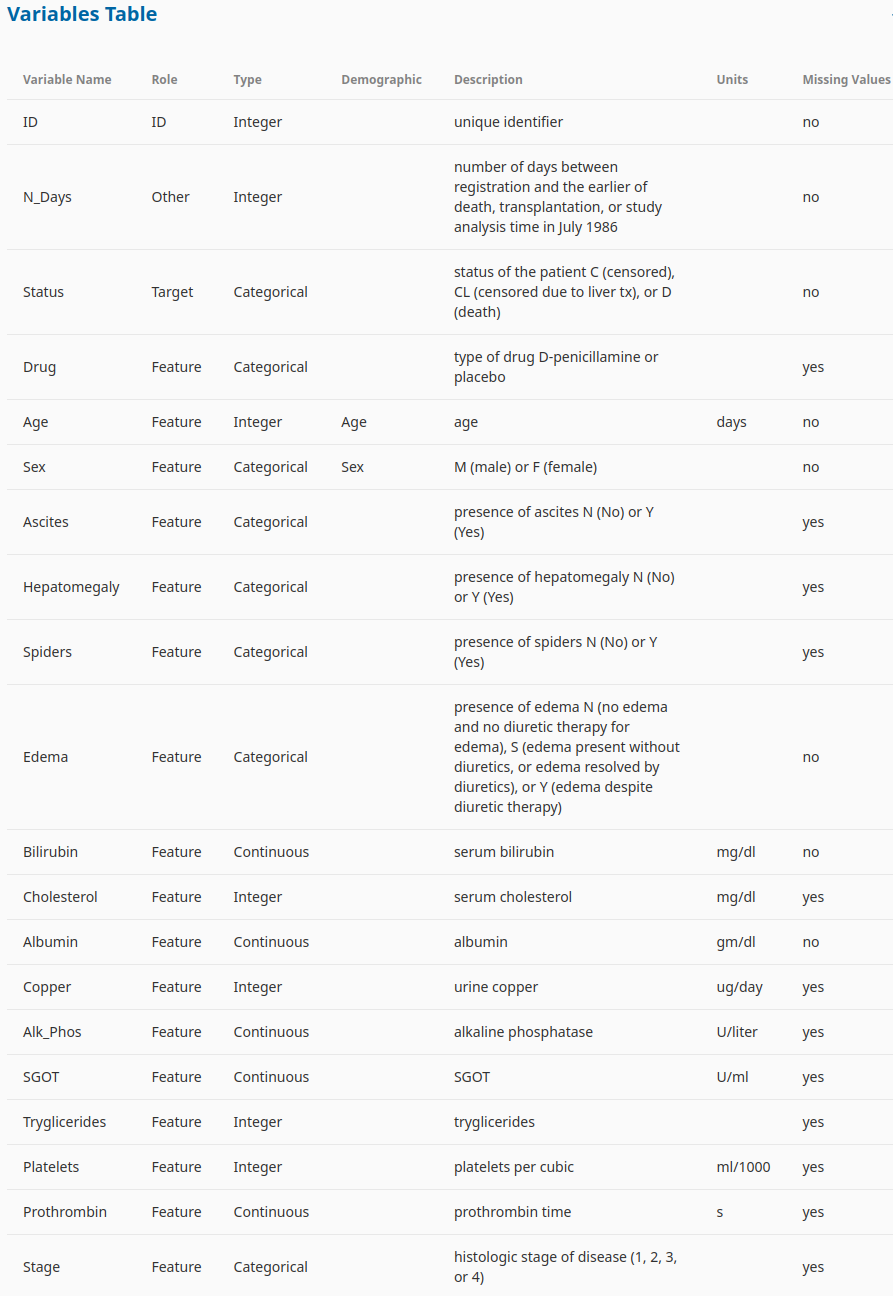
Cada repositorio creado debe tener un archivo “README” con la información del archivo, así como el archivo Jupyter-Notebook, el dataset a utilizar y este archivo de instrucciones. Dentro del README se debe detallar en qué consiste el análisis, qué pasos se siguen, qué pruebas se abordaron y qué modelos se utilizaron.

Como sugerencia, hay que ir tomando capturas de las gráficas, de los procesos más importantes y de las conclusiones, serán útiles para más adelante, hay que guardarlas en el mismo repositorio en una carpeta llamada “Screenshots” o similar.

Para entregar el proyecto, se deben enviar los siguientes archivos/datos:

* El archivo de Jupyter, el dataset y todo lo necesario para ejecutarlo.
* Este archivo de instrucciones con el nombre del alumno, número de control y número de proyecto.
* El enlace al repositorio público en GitHub del proyecto.

**Figura 1.** Información sobre el dataset.



Fuente: JOAKIM ARVIDSSON en Kaggle (2023).

Métrica de evaluación:

El proyecto tiene un valor de 40 puntos sobre la calificación final de la unidad, el puntaje estará dado por:

|  |  |  |
| --- | --- | --- |
| Criterio | Valor | Métrica |
| ETAPA 1. | 10 puntos | Excelente: 10 puntos (todos los puntos realizados correctamente).  Bien: 8 puntos (máximo 2 puntos no realizados totalmente bien o incompletos).  Regular: 5 puntos (hasta 3 puntos incompletos o con errores).  No aceptable: 0 puntos (más de 3 puntos incorrectos, incompletos o con errores)  Anulación si se detecta plagio. |
| ETAPA 2. | 10 puntos. | Excelente: 10 puntos (todos los puntos realizados correctamente).  Bien: 8 puntos (máximo 2 puntos no realizados totalmente bien o incompletos).  Regular: 5 puntos (hasta 4 puntos incompletos o con errores).  No aceptable: 0 puntos (más de 4 puntos incorrectos, incompletos o con errores)  Anulación si se detecta plagio. |
| ETAPA 3. | 10 puntos. | Excelente: 10 puntos (todos los puntos realizados correctamente).  Bien: 8 puntos (máximo 2 puntos no realizados totalmente bien o incompletos).  Regular: 5 puntos (hasta 4 puntos incompletos o con errores).  No aceptable: 0 puntos (más de 4 puntos incorrectos, incompletos o con errores)  Anulación si se detecta plagio. |
| ETAPA 4. | 10 puntos. | Excelente: 10 puntos (todos los puntos realizados correctamente).  Bien: 5 puntos (visualización incompleta o carga al repositorio incorrecta).  No aceptable: 0 puntos (no hay visualización ni carga en repositorio)  Anulación si se detecta plagio. |