Proyecto: EDA, Regresión, Hipótesis

instituto tecnológico de morelia

Proyecto.

M.I.A. WILFRIDO CORTES OROZCO

2023

|  |  |  |
| --- | --- | --- |
| Nombre del Alumno | Número de control | Número de Proyecto |
| Alan Jafet Leon Loeza | 20120126 | 2 |

# Descripción y objetivos del proyecto.

Objetivos:

* Que el estudiante pueda utilizar Git para llevar a cabo el control de versiones.
* Que el estudiante pueda realizar el análisis exploratorio de los datos completo sobre un dataset real.
* Que el estudiante ponga en práctica sus conocimientos teóricos y ejercite su capacidad de análisis.
* Que el estudiante demuestre sus habilidades al crear un modelo de regresión lineal multivariable.
* Que el estudiante se familiarice con las pruebas de hipótesis.

Descripción de la actividad:

Este proyecto consiste en cinco etapas principales:

1. Exploración de los datos.
2. Manipulación y tratamiento de los datos.
3. Construcción de un modelo de regresión lineal.
4. Hipótesis.
5. Publicación del proyecto.

Las instrucciones 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 “EDA\_LINEAR-REGRESSION\_HYPOTHESIS\_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.

Métrica de evaluación:

El proyecto tiene un valor de 30 puntos sobre la calificación final de la unidad, el puntaje estará dado por:

|  |  |  |
| --- | --- | --- |
| Criterio | Valor | Métrica |
| ETAPA 1 y ETAPA 2. | 12 puntos | Excelente: 12 puntos (todos los puntos realizados correctamente).  Bien: 8 puntos (máximo 2 puntos no realizados totalmente bien o incompletos).  Regular: 6 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 y ETAPA 4. | 12 puntos. | Excelente: 12 puntos (todos los puntos realizados correctamente).  Bien: 8 puntos (máximo 2 puntos no realizados totalmente bien o incompletos).  Regular: 6 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 5. | 6 puntos. | Completo: 6 puntos. El proyecto se encuentra público en su respectivo repositorio en GitHub.  Incompleto: 0 puntos. El proyecto no se encuentra público en su respectivo repositorio en GitHub. |