UNIDAD 4 – PROYECTO: ANÁLISIS DE SENTIMIENTOS

instituto tecnológico de morelia

Proyecto – UNIDAD 4.

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 aplicar técnicas de procesamiento de lenguaje natural para extraer información de textos.
* Que el estudiante pueda combinar los conocimientos de ciencia de datos que tiene con el procesamiento de lenguaje natural.
* Que el estudiante pueda realizar predicciones incluso en un dataset que solo contiene datos de tipo texto.

Descripción de la actividad:

Este proyecto consiste en cuatro etapas principales:

1. EDA.
2. Creación de un analizador de sentimientos.
3. Creación de modelos de ML para realizar predicciones.
4. Presentación del proyecto.

En esta ocasión se utilizarán 1 conjuntos de datos llamado “reviews.csv”, este dataset contiene información sobre las reseñas de personas aleatorias respecto a la adaptación a live-action del anime “One Piece”.

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 “PLN\_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.

Métrica de evaluación:

El proyecto tiene un valor de 60 puntos sobre la calificación final de la unidad, además existe la posibilidad de obtener puntos adicionales.

El puntaje estará dado por:

|  |  |  |
| --- | --- | --- |
| Criterio | Valor | Métrica |
| ETAPA 1. | 10 puntos | Excelente: 10 puntos (todos los pasos realizados correctamente).  Bien: 8 puntos (máximo 2 pasos no realizados totalmente bien o incompletos).  Regular: 5 puntos (hasta 3 pasos incompletos o con errores).  No aceptable: 0 puntos (más de 3 pasos incorrectos, incompletos o con errores)  Anulación si se detecta plagio. |
| ETAPA 2. | 20 puntos. | Excelente: 20 puntos (todos los pasos realizados correctamente).  Bien: 15 puntos (máximo 2 pasos no realizados totalmente bien o incompletos).  Regular: 10 puntos (hasta 4 pasos incompletos o con errores).  No aceptable: 0 puntos (más de 4 pasos incorrectos, incompletos o con errores)  Anulación si se detecta plagio. |
| ETAPA 3. | 20 puntos.  (+10 puntos adicionales posibles) | Brutal: 30 puntos (10 puntos adicionales si se realizan todos los pasos correctamente y, además, se emplean redes neuronales a través de TensorFlow/Keras).  Excelente: 20 puntos (todos los pasos realizados correctamente).  Bien: 15 puntos (máximo 2 pasos no realizados totalmente bien o incompletos).  Regular: 10 puntos (hasta 4 pasos incompletos o con errores).  No aceptable: 0 puntos (más de 4 pasos incorrectos, incompletos o con errores)  Anulación si se detecta plagio. |
| ETAPA 4. | 10 puntos. | Excelente: 10 puntos (todos los pasos realizados correctamente).  No aceptable: 0 puntos (no hay carga en el repositorio)  Anulación si se detecta plagio. |