



## INFORME FINAL DE AVANCE PROYECTO APLICADO II

### MAESTRÍA EN CIENCIA DE DATOS

**FECHA:** 9 de Diciembre de 2023

#### 1. Datos generales del Proyecto Aplicado

Título del Proyecto: Análisis comparativo de Modelos de Segmentación en imágenes de tomografía computarizada (CT) del área abdominal
Nombre del/los Estudiante(s): Jan Polanco Velasco y Stefania Astudillo Bello
Director del Proyecto Aplicado: Dr. Julián Gil González
Codirector/ Asesores (si aplica): NO APLICA
Grupo de Investigación o Empresa que apoya (si aplica): NO APLICA
Fecha de Inicio del proyecto: 18 de Septiembre de 2023

#### 2. Resumen del Trabajo de Grado

##### Objetivo General:

Realizar un análisis comparativo de modelos de segmentación en imágenes de tomografía computarizada (CT) del área abdominal, con el fin de determinar el modelo más efectivo para obtener una segmentación precisa y confiable, que logre los mejores resultados en términos de precisión, recall y otras métricas relevantes.

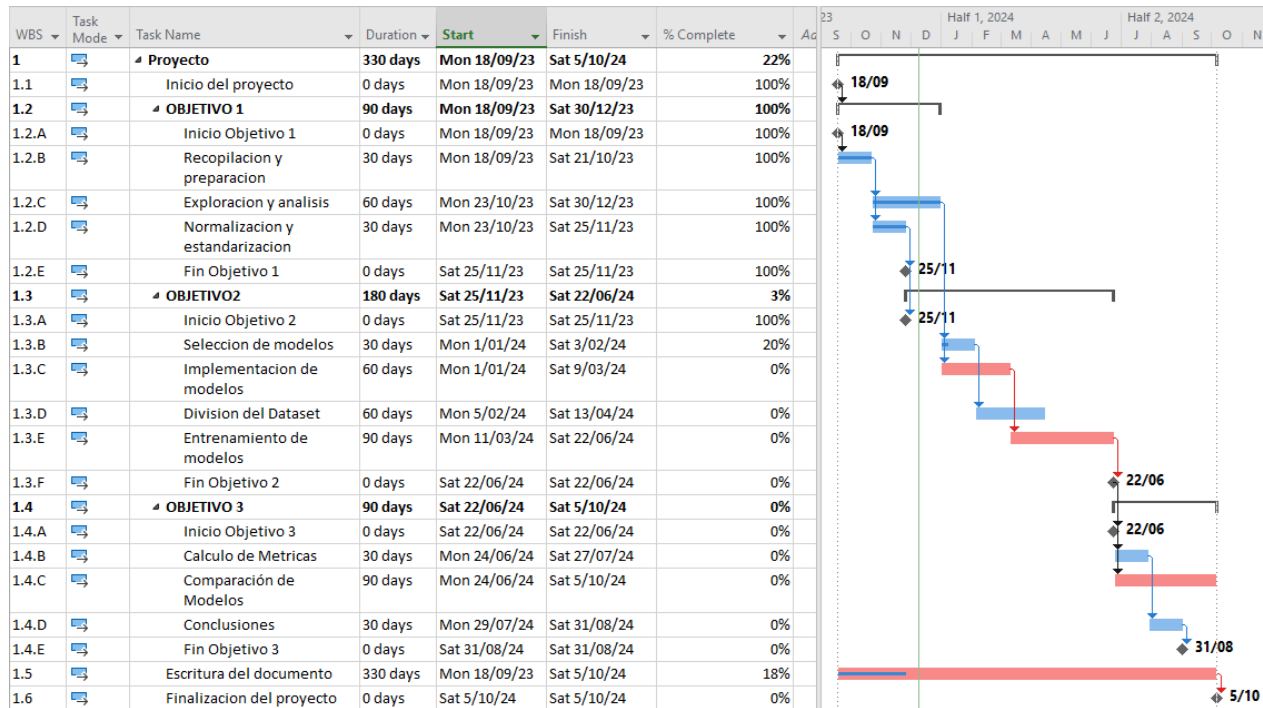
##### Objetivos Específicos:

- Realizar un procesamiento previo y exhaustivo de las imágenes de tomografía computarizada (CT) del área abdominal, con el objetivo de eliminar el ruido y mejorar la calidad de las imágenes, facilitando así el proceso de segmentación.
- Implementar y entrenar tres modelos definidos de *Deep Learning* para la segmentación de las imágenes de tomografía computarizada (CT) del área abdominal, utilizando modelos adecuados y técnicas avanzadas de *Machine Learning*, con el propósito de lograr una segmentación precisa y confiable.



- Realizar la validación del modelo de segmentación obtenido, evaluando su desempeño en términos de precisión, *recall* y otras métricas relevantes.

#### Cronograma de Actividades propuesto (\*):



#### Frecuencia de Reuniones con el Director:

1 Reunión mensual.

### 3. Resultados, avances y estado del Proyecto Aplicado

Hemos logrado completar todas las tareas establecidas con el objetivo 1. Sin embargo, es importante mencionar que nos enfrentamos a un leve retraso en el cronograma, específicamente en lo que concierne al manejo y procesamiento del dataset seleccionado. Este desafío surgió debido a la ausencia de metadatos adecuados asociados a las máscaras del dataset proporcionado.

Para abordar este inconveniente, exploramos diversas alternativas, incluyendo datasets provenientes de otras competencias. No obstante, dado que se trata de información médica sensible, el acceso a estas imágenes y máscaras implicó tiempos de espera prolongados, lo cual ha contribuido al mencionado retraso.

Implementamos medidas correctivas buscando soluciones efectivas para mitigar estos contratiempos y avanzar de manera eficiente. A continuación detallamos los primeros resultados obtenidos

- **Análisis y Preprocesamiento del Dataset:** Hemos realizado una evaluación exhaustiva y un preprocesamiento meticuloso del conjunto de datos. Este paso es crucial para asegurar la calidad y la relevancia de los datos para nuestros análisis posteriores.
- **Caracterización del Dataset:** Se ha llevado a cabo una caracterización detallada del dataset, lo que ha permitido un entendimiento profundo de sus atributos y particularidades. Este entendimiento es fundamental para guiar nuestras estrategias de modelado y análisis.
- **Identificación de Ruido en el Dataset Inicial:** Hemos identificado y abordado con éxito el ruido presente en el dataset, particularmente aquel asociado a la camilla utilizada para posicionar a los pacientes. Este ruido, si no se maneja adecuadamente, podría haber comprometido la integridad de nuestros resultados.
- **Elaboración y Optimización de Código:** Se ha desarrollado y optimizado código específico para este proyecto, lo que ha resultado en una versión mejorada y más eficiente del mismo, adecuada para manejar las complejidades y el tamaño de nuestro dataset.
- **Aumentación de Datos:** Dada la envergadura del dataset, se ha implementado una estrategia de aumentación de datos. Esto implica la generación de información sintética adicional para enriquecer nuestro conjunto de datos, asegurando así una base de datos más robusta para el entrenamiento de nuestros modelos.

Estos avances representan pasos significativos hacia el logro de nuestros objetivos de proyecto aplicado. Toda esta información, junto con detalles adicionales y análisis, se encuentra documentada de manera exhaustiva en el informe adjunto y en el repositorio de Github.

Porcentaje de avance de cada objetivo específico:

El porcentaje de avance es del 22% de acuerdo a lo reportado en el documento de Project



Pontificia Universidad  
**JAVERIANA**  
Cali

Firmas del/los estudiante(s), director del Proyecto y codirector si aplica:

Jan Polanco Velasco

Dr. Julián Gil González

Stefania Astudillo Bello