



# Tecnológico de Monterrey

Inteligencia artificial avanzada para la ciencia de datos 2

Gpo 501

## **Docentes**

Dr. Benjamín Valdés Aguirre

Ma. Eduardo Daniel Juárez Pineda

Dr. Ismael Solis Moreno

Dr. José Antonio Cantoral Ceballos

Dr. Carlos Alberto Dorantes Dosamantes

## **Integrantes**

Carlos Rodrigo Salguero Alcántara	A00833341
Diego Perdomo Salcedo	A01709150
Dafne Fernández Hernández	A01369230
José Emiliano Riosmena Castañón	A01704245
Luis Arturo Rendón Iñarritu	A01703572

Querétaro, Querétaro

## 1.0 Preparación de los datos

Con base a los resultados obtenidos en la iteración anterior, consideramos que debíamos realizar cambios a los datos para lograr obtener mejores resultados del modelo. En este documento, se hablará a detalle de la preparación de los nuevos datos para el desarrollo de una nueva versión del modelo.

### 1.1 Trasfondo

Durante la fase de modelado (modeling 3, 4 y 5), a pesar de los cambios en el tamaño del dataset, las arquitecturas utilizadas y los ajustes realizados a los parámetros de los modelos, no logramos obtener resultados satisfactorios con estos últimos. Por lo que fue necesario buscar otro enfoque para tratar de obtener un desempeño decente. Como solución, se cambió el etiquetado del dataset, pasando de ser clasificación a segmentación, utilizando Bounding - Boxes para marcar la ubicación exacta de cada vaca en la imagen por medio de coordenadas.

## 2.0 Limpieza de los datos

Para el proceso de limpieza de los datos hemos cambiado la forma del etiquetado de las imágenes, cambiando de clasificación a segmentación, utilizando las 8,000 imágenes originales del dataset.

### 2.1 Segmentación

Se reclasificaron las imágenes cambiando del conteo de número de vacas por imagen, a marcar con recuadros, en dónde se encuentran las vacas en cada imagen. En la preparación original de los datos, colaboramos con otro equipo para realizar el etiquetado, ambos equipos etiquetando para clasificación y segmentación, teniendo como resultado dos Datasets disponibles para ambos equipos. En las iteraciones pasadas, estábamos utilizando el primer dataset (clasificación), pero los resultados de los modelos no fueron eficientes, por lo que cambiamos al segundo dataset (segmentación).

El proceso de segmentación se llevó a cabo utilizando Roboflow, una herramienta especializada para el etiquetado y preparación de datasets en visión por computadora. El objetivo principal fue segmentar las imágenes, por medio de Bounding - Boxes, encerrando en

un recuadro a cada vaca presente en la imagen de la fila de espera para ordeña en un momento específico del día.