



# Tecnológico de Monterrey

**Campus Querétaro**

Introducción General del Modelado en la Etapa de Modelado

Gamaliel Marines Olvera	A01708746
Uri Jared Gopar Morales	A01709413
José Antonio Miranda Baños	A01611795
María Fernanda Moreno Gómez	A01708653
Oskar Adolfo Villa López	A01275287
Luis Ángel Cruz García	A01736345

Inteligencia artificial avanzada para la ciencia de datos II  
Grupo 501

## Introducción

En la fase de modelado realizamos dos etapas para verificar que nuestro modelo cumpla con el objetivo de minería de datos a través de los criterios de éxito:

1. El modelo identifica con un 80% de precisión el estado de cada una de las camas.
2. El modelo entrega el análisis de todas las camas en el 80% de las imágenes a lo largo del día.

Para ello, en la primera etapa se decidía si se debía pasar a una iteración desde un momento temprano y en la segunda etapa ya se buscaba cumplir eficientemente con las métricas que fueran necesarias para justificar el primer criterio de éxito. Una vez validado, se da paso a validar el segundo criterio. Además de los índices, se implementaron gráficas que mostrarán las predicciones en elementos específicos. Cabe destacar que el análisis detallado de cada uno de los índices y gráficas se describen en las diferentes iteraciones de modelado.

## Índices de modelado

En el siguiente diagrama se muestra el flujo de trabajo que se siguió para evaluar si se pasaba a la siguiente etapa de los índices o si se requería una nueva iteración.

Índices de modelado	Validación de índice	Criterio de éxito validado
Etapa 1		
Accuracy	Mayor al 95%	1
Matriz de confusión	Validación visual	1
Etapa 2		
Gráfica de distribución de predicciones por cama	Validación visual	1
Etapa 3		
F1 score	Mayor al 90%	1
Average Accuracy per Bed	Mayor al 80%	1
Etapa 4		

Distribución horaria del accuracy	Validación visual (Accuracy mayor al 80% en más del 80% del día)	2
-----------------------------------	--	---

En donde se puede notar que en caso de no contar con la validación del índice no se utilizaban los de la siguiente etapa. Además, se puede notar que el segundo criterio se validó al final, debido a que era menos prioritario y a que deriva del primero.

Los índices se describen a continuación:

$$Accuracy = \frac{Correct\ predictions}{Total\ predictions} * 100$$

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

En donde:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

$$Average\ Accuracy\ per\ Bed = \frac{Total\ Matches\ of\ Beds\ A,\ B,\ and\ C}{Total\ Actual\ Beds\ A,\ B,\ and\ C}$$

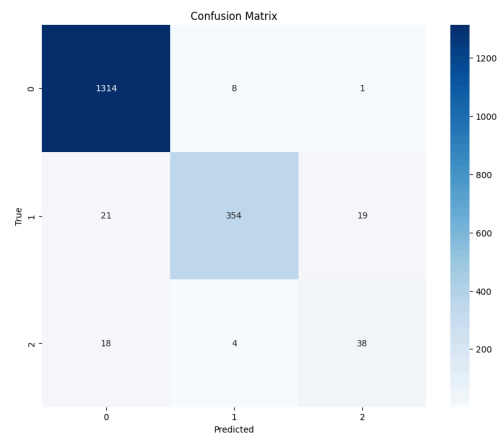
Siendo el Average Accuracy per Bed, el índice que esperamos tener con un valor mayor al 80% el cual es mencionado en el primer criterio de éxito.

## Gráficas utilizadas

Por otra parte, también utilizamos gráficas que nos ayudarán a verificar que nuestro modelo efectivamente cumplía con los criterios, a continuación se muestra su descripción e imágenes obtenidas en la última iteración a modo de ejemplo:

## Matriz de confusión

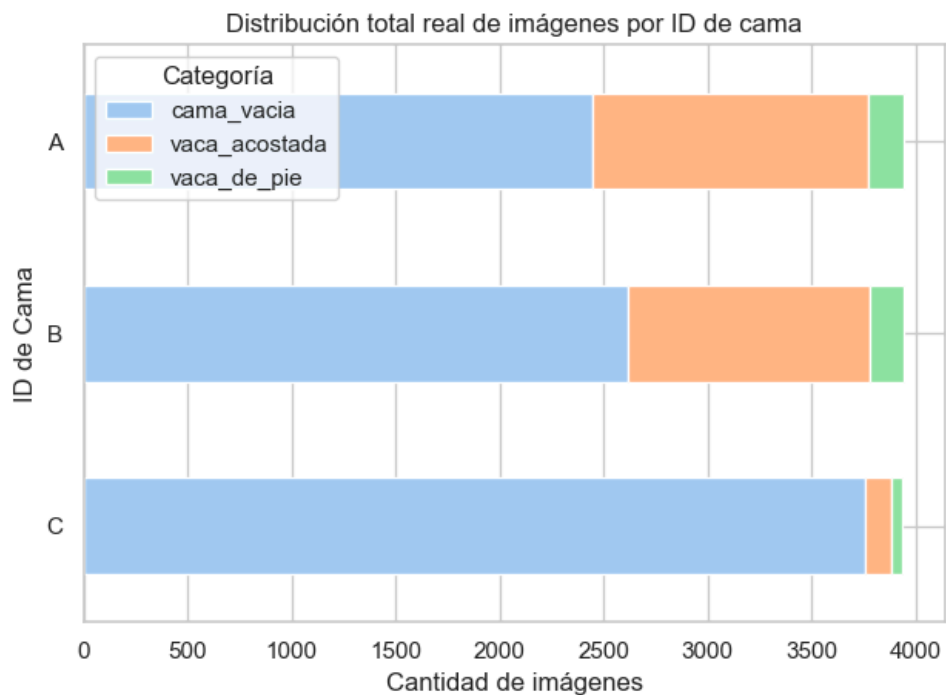
Es una gráfica que organiza las predicciones de un modelo en comparación con las clases reales (vaca acostada, vaca de pie, cama vacía), proporcionando una visión detallada de su desempeño.



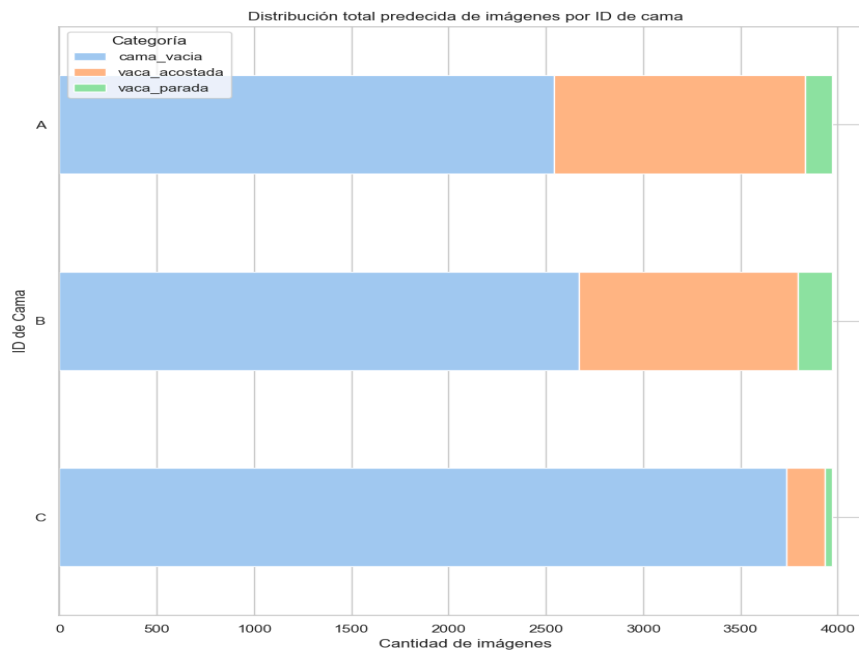
**Imagen 1:** Matriz de confusión del tercer modelo. (0: Cama vacía, 1: Vaca acostada, 2: Vaca de pie)

### Gráfica de distribución de predicciones por cama

Es una gráfica en la que podemos observar si las predicciones por cada cama concuerdan con los valores reales de cada una de ellas.



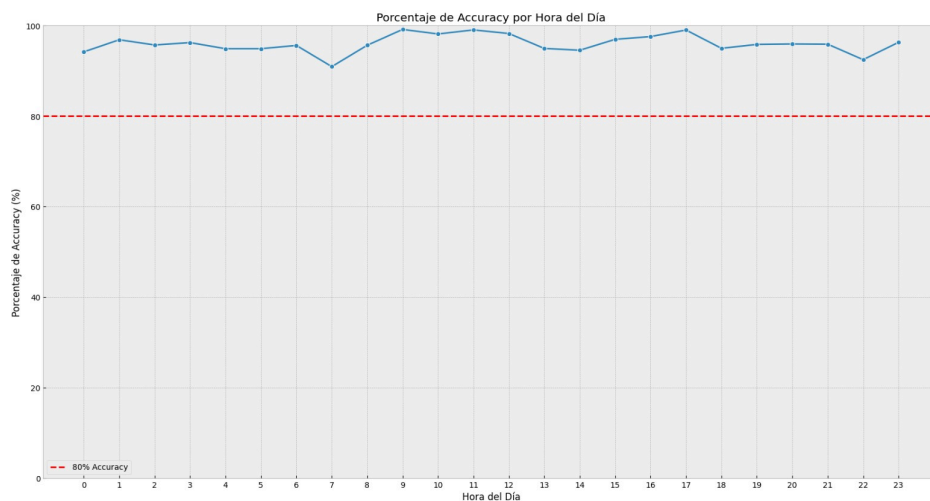
**Imagen 2:** Distribución de las imágenes de cada cama de acuerdo a las categorías. Dataset de entrenamiento. Clasificación manual.



**Imagen 3:** Distribución de las imágenes de cada cama de acuerdo a las categorías. Dataset de entrenamiento. Predicción del modelo.

### Distribución horaria del accuracy

En la gráfica se puede observar el accuracy en las diferentes horas del día, además, la línea horizontal muestra el accuracy menor aceptado, que es del 80%.



**Imagen 4:** Accuracy promedio a lo largo del día.