



Tecnológico de Monterrey

Campus Querétaro

Entendimiento de negocio

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

Resumen de negocio

El CAETEC (Campo Agro Experimental del Tec de Monterrey) ubicado en el KM 188 de la carretera México-Querétaro, funge como laboratorio para el aprendizaje de actividades agropecuarias debido a su tamaño (70 hectáreas) y sus condiciones.

En el área pecuaria, se cuenta con un establo con 6 corrales destinados a las 160 vacas que habitan esta área. En cada corral, las vacas tienen espacios para tomar el sol, descansar, comer y ser ordeñadas en una línea de ordeño que las ordeña de manera automática por medio de un brazo robótico que se encuentra cada 2 corrales.

Medir la cantidad y calidad de la leche es algo muy importante para el CAETEC, pues dependiendo de estos indicadores, es el precio que se le puede poner la leche. Por ende, siempre se busca una mayor y mejor producción de leche, este siendo el principal KPI que se tiene en el área pecuaria.

Una mejor producción de leche está relacionada con el descanso de la vaca, pues es en el momento que realizan la rumia. Como todos los animales, la digestión es un proceso en el que el individuo separan y absorben los nutrientes de los alimentos, y, debido a la anatomía de la vaca y su proceso fisiológico de la digestión, la vaca necesita tragar, regurgitar, volver a masticar y volver a tragar la comida, este proceso se llama rumia.

Este proceso puede tardar hasta 10 horas, dependiendo de qué tanta fibra tenga el alimento, así como una velocidad de tránsito y absorción ruminal de hasta 50 horas (dependiendo de qué tan digerible es el alimento). Es por eso que la rumia depende de la calidad del alimento (a mayor calidad, menor tiempo de rumia) y de otros factores como el celo de la hembra (si está en celo, disminuye la rumia) y el dolor y estrés. La posición ideal de rumia es en decúbito esternal (aunque también se puede hacer parado o caminando, pero es más largo este periodo al estar en estas posiciones al gastar energía en otras funciones extras a la rumia), es por eso que se prioriza una rumia de la vaca mientras esta está acostada (en un lugar cómodo para acostarse) que parada o caminando, ya que el gasto de energía adicional repercute negativamente en la producción de leche. La energía es relevante en la producción de leche, ya que cuando existen pérdidas grandes de energía, se presentan disfunciones del metabolismo que afectan radicalmente la producción de leche, pues la energía es esencial para el aprovechamiento de los nutrientes del alimento como lo son las proteínas.

Con eso en mente, el CAETEC llega con nosotros con la finalidad de ayudarlos para entender y mejorar su KPI de producción de litros de leche por hectárea, por medio de dos cámaras web (una en el área de las camas que es con la que vamos a

trabajar) que toma fotografías cada 5 minutos a 3 espacios de camas de las vacas todos los días del año para que se analicen las imágenes y se logre obtener algún descubrimiento que ayude a la mejora de la producción de leche, enfocada en el proceso de rumia de las vacas.

Objetivo de negocio

Identificar los patrones del estado de ocupación de las camas en tiempo real.

Criterios de éxito del objetivo de negocio

- ☐ Se obtiene conocimiento de los estados de las camas (vacía/vaca acostada/vaca parada) en tiempo real, es decir, cada que llegue una imagen al dataset.

Memorandum of understanding

El equipo de trabajo se compromete a cumplir el objetivo de minería de datos, medible mediante los criterios de aceptación estipulados en este documento. Es posible que durante el desarrollo del proyecto, surjan limitaciones o contingencias que requieran cambios en el objetivo o los criterios de aceptación, los cuales se harán teniendo siempre en cuenta el objetivo de negocio del socio formador.

El socio formador se compromete a proveer los datos necesarios para el correcto desarrollo del proyecto, los cuales incluyen, pero no se limitan a:

- Imágenes para la extracción de datos.
- Información respecto a sus necesidades y expectativas.
- Comunicación posterior para aclaración de dudas.

Mapeo de recursos

[Versión Gráfica](#)

Proporcionados por el socio formador:

- Dataset de 9634 imágenes en formato .jpg y de tamaño de 1920 x 1080px. [Link](#)
- Cámara web marca Logitech que tiene un alcance de 3 camas completas de las vacas.
- Asesorías con el Dr. Ivo Ayala. (Con una petición previa).

Adicionales:

- Asesorías con los profesores (Con una petición previa).

- Clases teóricas y prácticas.
- Laboratorio de AWS.
- 4 laptops con tarjetas gráficas:
 - RTX 4070
 - RTX 4080
 - GTX 1060
 - GTX 1650
- Otro equipo (No Name) con el cual podemos colaborar.
- Otros equipos con los cuales podemos pedir apoyo.
- Internet y sus recursos.
- Biblioteca Tec. [Link](#)

Riesgos y contingencias

Los diferentes riesgos se plasmarán en una matriz de riesgos, para ello, primero es importante definir las siguientes clasificaciones y métricas.

Efecto	Clasificación del impacto
Más de 2 semanas de retraso en el proyecto. Objetivo principal no alcanzado. Es de gran importancia y requiere comunicarse con el Socio Formador.	5
Retrasa los resultados del proyecto de 1 a 2 semanas. Objetivo no alcanzado de manera plena.	4
Retrasos internos de 3 días en los entregables de una fase. Los objetivos de más de una tarea no se logran.	3
Afecta a otras tareas en 1 - 3 días. Un objetivo de tarea no alcanzado.	2
Retraso de una actividad o función no esencial para el objetivo principal, sin impacto.	1

Tabla 1. Clasificación por impacto en el cronograma.

Efecto	Factor de probabilidad de riesgo
$80\% \leq \text{Probabilidad de riesgo} \leq 100\%$	5
$60\% \leq \text{Probabilidad de riesgo} < 80\%$	4

40% ≤ Probabilidad de riesgo < 60%	3
20% ≤ Probabilidad de riesgo < 40%	2
0% < Probabilidad de riesgo < 20%	1

Tabla 2. Clasificación de probabilidad de riesgo.

Nivel de riesgo	Acciones
1-6	No hay que tomar ninguna medida.
7-12	Se deberán discutir medidas como trabajo en horarios más extensos, colaboración presencial.
13-18	Se deberá discutir con el equipo para tomar mayor atención en las actividades que tengan un mayor impacto para el alcance del objetivo, también, se deberán discutir medidas como trabajo en horarios más extensos y con colaboración presencial.
19-25	Se entrará en comunicación con profesores y Socio Formador, se buscará hacer una entrega, que si bien puede que no cumpla con el objetivo, sí, pueda ser mejorable.

Tabla 3. Niveles de riesgo.

Donde el nivel de riesgo es igual al producto de la probabilidad por el impacto de cada riesgo.

Número	Riesgo	Probabilidad	Impacto	Nivel de riesgo	Estrategia de mitigación	Estrategia de contingencia
1	Contar con una gran cantidad de datos que no brinden información relevante para lograr el objetivo del	3	4	12	Ser detallado en la fase de exploración y preprocesamiento de los datos para filtrar y seleccionar aquellos que sean más relevantes para los objetivos del proyecto.	Búsqueda de alternativas como el uso de datos similares en internet o el uso de transformaciones en las imágenes con las que se cuentan. También está la posibilidad de pedir al Socio Formador que obtenga más imágenes, pero

	proyecto.					deberá ser discutido dependiendo del tiempo restante.
2	Complejidad en el desarrollo del algoritmo mayor al previsto, lo que resulta en un resultado ineficiente para el Socio Formador.	3	4	12	Dividir el desarrollo en fases con hitos de validación temprana, para identificar problemas de complejidad y ajustar el enfoque técnico según sea necesario.	Contar con propuestas de solución alternas, como el uso de otros modelos o de técnicas de afinación de parámetros, que en caso de ya no ser implementadas debido al tiempo, cuenten con una mención dentro de la documentación para poder ser aplicados a futuro.
3	Mal entendimiento de la necesidad del Socio Formador.	2	5	10	Realizar reuniones frecuentes con el Socio Formador para aclarar los objetivos y expectativas del proyecto, y documentar claramente los requisitos clave, además, ser claro y efectivo al momento de tener comunicación.	Buscar apoyo para la definición de los objetivos del Socio Formador con todos los integrantes e inclusive con otros equipos, además, definir los objetivos o cualquier información que se quiera validar aunque no se tenga claro si es correcta, para de esta manera buscar la aprobación del SF o en su caso recibir retroalimentación para correcciones.

4	El modelo no generaliza bien a nuevos datos, lo que resulta en un desempeño o pobre en entornos reales.	3	3	9	División de los datos en conjuntos de entrenamiento, validación y prueba, variando tanto en cantidad como en datos en sí, buscando complementar con datos de entrenamiento más variados. Utilizar pruebas de validación cruzada en diferentes etapas del desarrollo del modelo.	Hacer correcciones en los datos utilizados, implementando técnicas como transformaciones, además, estar consciente de la distribución de patrones para poder probar con diferentes pesos en el modelaje. En caso de que ya no sea factible debido al tiempo, se deberá comunicar al SF que el funcionamiento contará con mejores resultados en condiciones específicas.
5	Cambios en la infraestructura del Socio Formador destinada al proyecto.	2	2	4	Mantener una comunicación constante para anticipar cambios y adaptar el proyecto, además de realizar el proyecto lo más estandarizado posible para ser compatible con diferentes tecnologías y con diferentes especificaciones.	Implementar tecnologías que permitan el uso de cómputo en la nube, para que de esta manera no se dependa tanto de las características de la infraestructura.
6	Disponibilidad de los integrantes del equipo limitada debido a enfermedades, accidentes	1	3	3	No asignar muchas tareas a una sola persona y promover confianza entre los integrantes para comunicar efectivamente los percances.	Redistribución de tareas entre los miembros del equipo y buscar apoyo en otros equipos de manera responsable.

	o condicione s imprevista s.					
7	Disponibili dad limitada por parte del Socio Formador.	1	3	3	Ser eficiente y conciso en el tiempo que se tenga para hablar con el SF.	Utilizar medios remotos para poder obtener información concisa. En nuestro caso el principal medio será el correo electrónico.
8	Dificultad en la integració n del modelo en la infraestruc tura del Socio Formador.	2	1	2	Tener presente desde el inicio diferentes formas de poder implementar el modelo en la infraestructura señalada.	Contar con un manual de despliegue para poder compartir opciones de la integración de maneras y en infraestructuras que nosotros hayamos comprobado, además, considerar la implementación de tecnologías que permitan el uso de cómputo en la nube que faciliten la integración.

Tabla 4. Matriz de riesgos priorizados.

Determinar objetivos de Data Mining.

Data Mining Goal:

- Entregar información del estado de las vacas en las camas (vacía/acostada/parada), junto con la etiqueta del tiempo para cada una de las fotos que se reciben.

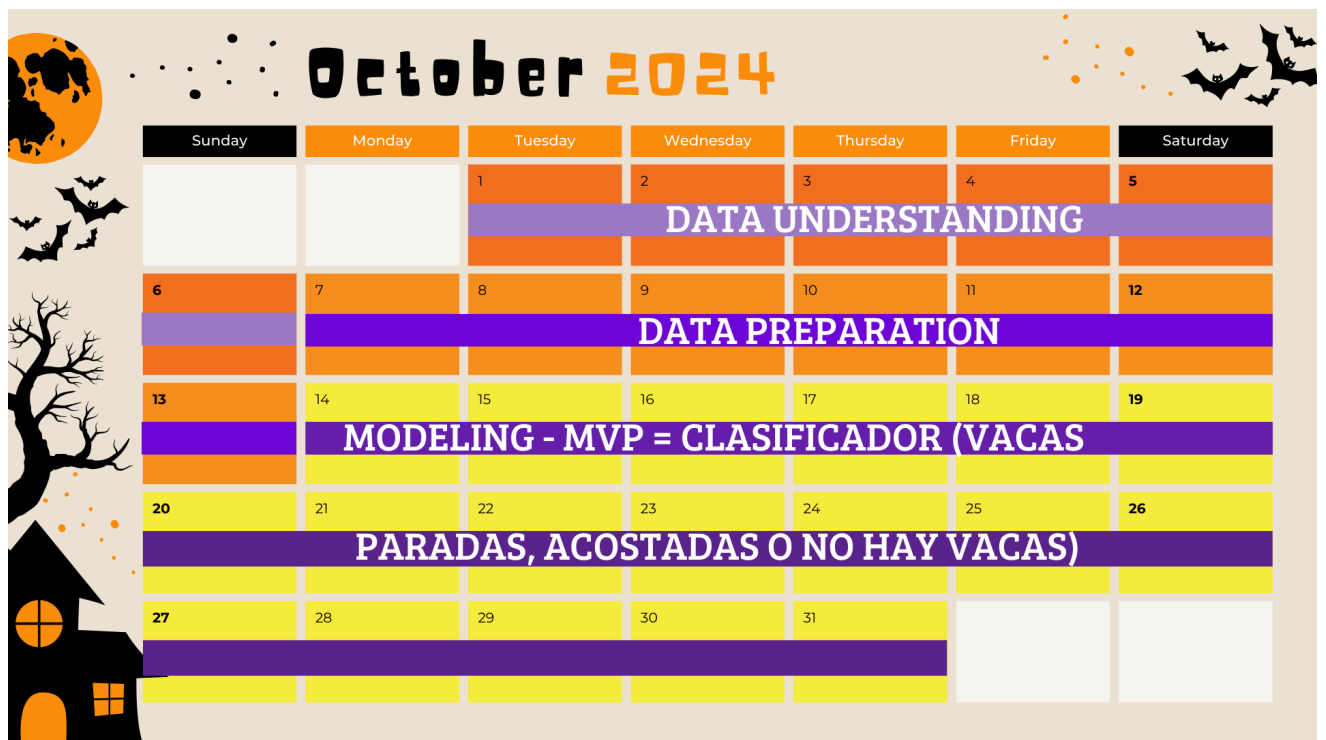
Data Mining Success Criteria.

- El modelo identifica con un 80% de precisión el estado de cada una de las camas. Esta decisión está basada en el funcionamiento de diferentes modelos en un contexto similar, donde existe un rango de entre el 73% al 96% de precisión, donde también existe variabilidad debido al tiempo de

respuesta de las predicciones (Gao et al., 2024), y nosotros al buscar un modelo que pueda obtener predicciones en máximo 5 minutos, establecimos un porcentaje de precisión alcanzable que no requieren muchos recursos para realizar las predicciones.

- El modelo entrega el análisis de todas las camas en el 80% de las imágenes a lo largo del día.

Plan inicial





En el mes de octubre del presente año, tendremos la primera semana del 1 - 6, nuestro primer análisis de los datos:

Para esta actividad nos juntaremos en equipo, esto con el propósito de recopilar la información que tengamos sobre la visita al CAETEC, de igual manera hablaremos sobre los artículos de investigación que hayamos leído.

Al realizar esta actividad lo que buscamos es que todos los integrantes del equipo nos encontremos en una misma sintonía, cuál es nuestro principal propósito para el proyecto y cuáles son nuestros recursos para cumplir dicho objetivo.

Para la semana 2 que abarca del día 7 - 13 de octubre:

Tenemos como objetivo lograr realizar nuestra preparación de los datos, lo lograremos al dividir el dataset con otro equipo. Lo que buscamos es poder clasificar las imágenes proporcionadas por el socio de la siguiente manera (hay vacas paradas, no hay vaca o la vaca está acostada). Al realizar esta actividad tendremos nuestro dataset listo en el menor tiempo posible, ya que somos 5 integrantes por equipo y las imágenes que vamos a clasificar son alrededor de 10,000. Esta actividad nos ayudará para poder entrenar nuestro modelo y que identifique la posición de las vacas en las fotografías.

Las últimas semanas del mes de octubre que es del 14 - 31:

Nuestro objetivo como equipo es tener un modelo funcionando el cual ya pueda

brindar la información de la cantidad de tiempo que las vacas están en las camas, cuál es la cama que más o menos se ocupa.

Terminología

- **CAETEC:** Campo Agro Experimental del Tec de Monterrey
- **KPI:** Clave de Desempeño (Key Performance Indicator)
- **Rumia:** Proceso digestivo de los rumiantes que involucra la regurgitación, remasticación y nueva deglución del alimento.
- **Decúbito esternal:** Posición en la que el animal está acostado sobre el esternón.
- **Metabolismo:** Conjunto de procesos bioquímicos que ocurren dentro de un organismo.
- **Validación cruzada (Cross-validation):** Técnica utilizada en el análisis de datos para evaluar el rendimiento de un modelo.
- **Dataset:** Conjunto de datos utilizado para entrenar y validar modelos de inteligencia artificial.
- **Preprocesamiento:** Etapa de limpieza y preparación de los datos antes de su análisis.
- **Ordeño automático:** Proceso automatizado en el que las vacas son ordeñadas mediante máquinas robóticas.
- **Tránsito ruminal:** Tiempo que tarda el alimento en ser digerido y absorbido en el rumen.
- **Brazo robótico:** Dispositivo mecánico usado para automatizar tareas como el ordeño.
- **Exploración de datos:** Fase inicial del análisis de datos para identificar patrones y características relevantes.
- **Minería de datos:** Proceso de analizar grandes conjuntos de datos para extraer información útil.

Referencias

CONtexto Ganadero. (2017, June 9). *El proceso de la rumia: Tránsito y absorción ruminal*. CONtexto Ganadero.
<https://www.contextoganadero.com/ganaderia-sostenible/el-proceso-de-la-rumia-tran-sito-y-absorcion-ruminal>

García Alegría, K. (2018, May 4). *Importancia de la energía en la dieta de los rumiantes*. Engormix.

https://www.engormix.com/lecheria/formulacion-ration-vaca-lechera/importancia-energia-dieta-rumiantes_a42115/

Hernández, J. (n.d.). *Comportamiento bovino de alimentación, rumia y bebida*. Producción Animal.

https://www.produccion-animal.com.ar/etologia_y_bienestar/etologia_bovinos/04-comportamiento_bovino_de_alimentacion_rumia_y_bebida.pdf

Japan International Cooperation Agency (JICA). (n.d.). *Sistema de alimentación en rumiantes*.

https://www.jica.go.jp/Resource/project/bolivia/3065022E0/04/pdf/4-3-1_07.pdf

TecScience. (n.d.). *CAETEC se convierte en un gran laboratorio agro*. TecScience.

<https://tecscience.tec.mx/es/tecnologia/caetec-laboratorio/>

U.S. Food & Drug Administration. (2021, September 20). *Cómo comen pasto las vacas*.

<https://www.fda.gov/animal-veterinary/animal-health-literacy/como-comen-pasto-las-vacas>

Gao, J., Bambrab, C. K., Parihar, N., Kshirsagar, S., Mallarapu, S., Yu, H., Wu, J., & Yang, Y. (2024). Analysis of Various Machine Learning Algorithms for Using Drone Images in Livestock Farms. *Agriculture*, 14(4), 522.

<https://doi.org/10.3390/agriculture14040522>