



Tecnológico de Monterrey

Campus Querétaro

Entrega

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

El presente documento presenta la fase de *Entrega* del modelo desarrollado, siguiendo el marco metodológico CRISP-DM (Cross-Industry Standard Process for Data Mining). En esta etapa, se busca implementar el modelo desarrollado en el entorno de producción de nuestro socio formador el CAETEC para que pueda ser utilizado de manera continua y efectiva, cumpliendo con los objetivos de negocio planteados inicialmente.

El principal objetivo de negocio de este proyecto es automatizar la clasificación de estados de ocupación en tiempo real, proporcionando resultados consistentes y accionables a partir de imágenes recibidas de manera continua. Para ello, el modelo de redes neuronales convolucionales (CNN) desarrollado ha sido optimizado e integrado en una arquitectura que permite procesar datos en tiempo real y generar predicciones accesibles a través de una interfaz.

Durante esta fase, se establecieron protocolos de monitoreo para garantizar el desempeño del modelo en producción, asegurando que la precisión lograda en el conjunto de prueba (96%) se mantenga estable en el entorno operativo. Asimismo, se implementaron mecanismos de retroalimentación para actualizar el modelo cuando sea necesario, adaptándose a posibles cambios en los datos o en las condiciones del entorno.

A continuación, se describen los pasos seguidos para la implementación del modelo, los desafíos enfrentados durante la integración, las pruebas realizadas para validar su desempeño en producción y las recomendaciones para asegurar su mantenimiento y escalabilidad futura. Este análisis busca proporcionar una visión clara de cómo el modelo será utilizado para generar valor continuo y facilitar decisiones estratégicas basadas en datos.

Plan de Entrega

En este apartado se explicará la forma de entrega de nuestros resultados, de la evaluación se crea una estrategia para entregar los resultados de minería de datos al negocio. Para encontrar nuestro plan de entrega ingrese al siguiente enlace:

- [Plan de entrega](#)

Planeación, Monitoreo y Mantenimiento

Es necesaria si los resultados formarán parte de la operación del negocio y su ambiente. Evita periodos largos innecesarios de uso incorrecto de los resultados. Se elabora un plan detallado del proceso de monitoreo considerando el tipo específico de entrega

- [Plan de Monitoreo y Mantenimiento](#)

Reporte final

Este reporte presenta los hallazgos obtenidos mediante la implementación de un modelo CNN para la detección de vacas en las camas, desarrollado siguiendo la metodología CRISP-DM. Dando a conocer las desviaciones a la que nos enfrentamos de nuestro plan inicial para concluir con nuestra solución, este reporte está dirigido para nuestro Socio Formador.

- [Reporte final](#)

Documentación de la Experiencia

A continuación se detalla la experiencia del equipo a lo largo del desarrollo del proyecto, incluyendo elementos como desafíos, aprendizajes y recomendaciones para futuros proyectos.

Además, se mostrará una retroalimentación obtenida al entrevistar a uno de los usuarios del proyecto.

Análisis de nuestra experiencia en el equipo TC.

Dentro de nuestra experiencia como equipo, identificamos puntos que nos ayudaron a que el proyecto continuara de manera exitosa, por ejemplo, una buena comunicación con todos los involucrados, el compromiso con el proyecto, el seguimiento de la retroalimentación brindada y el uso de herramientas para el manejo de versiones y que permitían el trabajo colaborativo.

Siendo la comunicación con el Socio Formador algo fundamental para poder entender la necesidad del negocio y poder entregar la solución que resuelva la problemática.

Además, se evitaron enfoques engañosos como la consideración de modelos extras que no resultan eficientes considerando el objetivo de negocio y la duración del proyecto, un ejemplo de esto fue el descarte de un modelo que seleccionara automáticamente las camas de las vacas, el cual fue reemplazado por una interfaz de usuario que garantiza la selección de las camas de una forma rápida y precisa.

Así mismo tuvimos la oportunidad de aprender a utilizar diferentes herramientas de análisis de datos como es Tableau, la cual es una herramienta poderosa para realizar gráficos que nos proporcionan la información necesaria para analizar los resultados de nuestro modelo.

Otros de los aprendizajes obtenidos fueron las técnicas para realizar diferentes tipos de análisis de datos, con el propósito de medir nuestros resultados obtenidos y poder comparar diferentes modelos realizados para tomar decisiones de cuál tiene un mejor desempeño.

Sin embargo, sí encontramos algunos desafíos. Unos de ellos fueron durante el proceso de prueba de nuestra arquitectura, los cuales estaban relacionados con las herramientas de hardware, específicamente las versiones de Raspberry Pi. Para garantizar que nuestra solución pudiera ejecutarse de manera eficiente, fue necesario encontrar una Raspberry Pi con arquitectura de 64 bits, ya que las versiones anteriores carecen de la capacidad suficiente para soportar la carga computacional requerida.

Otras de las limitaciones que enfrentamos fueron en términos de capacidad de cómputo, lo cual impactó directamente en el tiempo necesario para entrenar nuestras soluciones más robustas. Estas limitaciones resultaron en tiempos de procesamiento prolongados, lo que destacó la importancia de contar con hardware más potente para optimizar el desempeño y la eficiencia del desarrollo.

Por último, nos dimos cuenta que al momento de realizar el análisis y preparación de los datos, no identificamos métricas precisas desde un inicio, lo que resultó en que dichas métricas se fueran agregando durante el proceso de modelado en forma de etapas, lo cual no es completamente incorrecto pero sí puede resultar confuso. Sin dejar de lado que se debió de considerar con mayor importancia el impacto de pocas imágenes de una clase en específico al momento de modelar, abriendo la posibilidad de etiquetar más imágenes antes de pasar al modelado.

Así que para mejorar el proceso, se recomienda que en el momento de hacer el plan de recursos, se haga una investigación con la que podamos establecer un criterio sólido que nos diga si el hardware con el que se cuenta es suficiente, o en dado caso, saber hasta qué punto estamos limitados.

Por otra parte, también es necesaria una investigación más profunda acerca de qué métricas son útiles en el uso de modelos similares, para así comenzar la fase de modelado con métricas precisas ya definidas, garantizando que todos los modelos sean evaluados con todas las métricas.

Sin dejar de lado que ahora que se cuenta con la experiencia de lo que sucede al contar con pocas imágenes de una clase y que a su vez pueden ser confundidas fácilmente con otra clase, como fue nuestro caso con vacas paradas y camas vacías, sabemos que es mejor hacer un etiquetado mayor desde el entendimiento y preparación de datos.

Entrevista al usuario final

Se realizó una entrevista final con el Socio Formador antes de la presentación del proyecto, en la que se documentó su opinión del resultado considerando la comunicación que existió a lo largo de las diferentes fases. Esto con el fin de conocer si hubo algunos elementos por mejorar o si se necesita algún acompañamiento específico.

¿Están satisfechos?

Ivo Neftali: Nos expresó su admiración por el hecho de que este proyecto haya sido completamente diseñado y desarrollado por nosotros como equipo, partiendo de una solución basada en las necesidades específicas del negocio, sin que ningún profesor o fuente externa nos proporcionará una solución predeterminada. Destacó cómo, gracias a nuestros conocimientos, logramos implementar una propuesta efectiva con un rendimiento sobresaliente.

Está satisfecho con los resultados obtenidos, así como con la documentación que entregaremos, la cual incluye manuales y planes de monitoreo que garantizarán la continuidad y sostenibilidad del proyecto. También valoró la manera en que integramos las retroalimentaciones que nos proporcionó, analizándolas cuidadosamente para tomar las decisiones más adecuadas que solucionarán de manera óptima la problemática planteada. Finalmente, destacó que el rendimiento alcanzado por nuestro modelo en la predicción de posibles escenarios es muy buena y considera que refleja el esfuerzo y la calidad del trabajo realizado.

¿Qué pudo hacerse mejor?

Ivo Neftali: Destacó que lo que podría mejorar sería el acompañamiento más cercano durante todo el proceso de desarrollo del proyecto, ya que, como socio formador y no como profesor, no tuvo la oportunidad de observar directamente los errores o la dinámica de trabajo del equipo. Sin embargo, hizo hincapié en que, aunque no estuvo presente con nosotros durante todo el proceso, quedó impresionado por el compromiso y la dedicación que los miembros del equipo demostraron hacia el proyecto. Mencionó que, al visitar otros equipos, notaba la falta de integrantes o la falta de atención a las retroalimentaciones, pero al interactuar con nosotros, percibía una actitud positiva y un fuerte deseo de avanzar y cumplir con los objetivos.

¿Qué apoyo adicional necesitan?

Ivo Neftali: Nos comentó que, al finalizar la presentación y concluir el proyecto, le gustaría que pudiéramos dedicar un día para enseñarle cómo implementar la solución entregada en CAETEC. Además, solicitó que le proporcionemos un contacto con el cual pueda comunicarse en caso de que surjan dudas o requiera asistencia en el futuro.

Aportaciones y aprendizajes personales

Oskar

Business understanding. Participé en el planteamiento de objetivos de negocio, además de la redacción de la documentación.

Data understanding. Me encargué del diseño del modelo de almacenamiento y flujo de datos, seleccionando las herramientas a utilizar y la arquitectura planteada.

Data preparation. Participé en la clasificación manual de los datos. Desarrollé un script en conjunto con Antonio para poder hacer la clasificación de manera más rápida. Participé en la toma de decisiones del número de imágenes y el formato a utilizar.

Modeling. Participé en el análisis de datos de la fase 2, que llevó al descubrimiento de errores en el modelo. Participé en el diseño del modelo para la fase 3, y las pruebas en esta. En conjunto con Uri, realicé el diseño y la implementación del análisis de datos final. Desarrollé, en conjunto con Luis, la interfaz gráfica de usuario que permite realizar la selección de las camas.

Evaluation. Participé en el diseño de las tablas que utilizamos para poder comparar nuestros modelos, y en la toma de decisiones para el siguiente paso.

Deployment. Desarrollé las funciones que permiten realizar predicciones con el modelo integrando la selección de las camas, y guardar los resultados en un archivo CSV. Implementé el manejo de errores y logging para este código. Empaqueté el código en una librería y la publiqué para facilitar la instalación. Desarrollé el manual de usuario.

El aprendizaje más importante que me llevo es que es muy importante definir métricas adecuadas desde el inicio para poder evaluar correctamente los modelos.

Antonio

Business understanding. Contribuí a definir los objetivos de negocio y a redactar la documentación.

Data understanding. Me encargué la limpieza y clasificación de las imágenes, proponiendo criterios para generar las futuras clases. Planteé también las dimensiones que tendría la imagen recortada y propuse alargar la imagen mediante OpenCv para recuperar tanta información como fuese posible.

Data preparation. Junto a Oskar generamos un script para clasificar de forma manual las imágenes, este las recorta, alarga, les da un identificador para conocer de qué cama proviene la imagen y las agrega a una carpeta con su label.

Modeling. Me vi involucrado de lleno en el desarrollo de todos los modelos, realizando cambios y mejoras, desde agregar transformaciones hasta cambiar la estructura del modelo. Con la ayuda de Oskar y Uri detectamos problemas en la segunda iteración y actuamos a la brevedad. Apoye a Uri en la prueba de arquitectura, instalando el sistema operativo y varias dependencias necesarias. Creamos el repositorio donde alojamos el proyecto y con la ayuda de Oskar lo organizamos y le dimos mantenimiento.

Evaluación. Apoye en la documentación y en la creación de métricas para evaluar el desempeño de los modelos y evalúe los mismos.

Deployment. Me encargué del reporte final y genere el modelo que se entregará al socio para que instale junto a la paquetería que desarrolló Oskar.

Gamaliel Marines

Business Understanding. Aporté ideas y ayudé a definir los objetivos de negocio, minería de datos y las necesidades de nuestro cliente. También aporte a los criterios

para definir qué significaría tener un proyecto exitoso. Junto con Mafer, hicimos el mapa de recursos.

Data Understanding. Aparte de hacer un listado de las herramientas y técnicas utilizadas y sus razones, hice una investigación y explicación de las herramientas y técnicas no utilizadas (big data) acompañada de una argumentación de nuestra decisión de no utilizarlas.

Data Preparation. Mi papel en esta etapa consistió en la clasificación del subset que se me asignó (se dividió equitativamente el dataset original y se asignó a cada uno de los integrantes su porción del dataset). En el proceso de etiquetado me aseguré que el dataset cumpliera con las políticas de privacidad y seguridad de datos establecidas por el equipo y acordadas con nuestro socio formador (junto con mafer, definimos esas políticas y el proceso para mantener esas políticas en y durante nuestro proyecto).

Modeling. Me encargue del entrenamiento y mejora de la segunda versión del modelo. La mejora consistía en minimizar el tiempo de cómputo y reducción de consumo de recursos. Documenté el proceso y redacté los resultados obtenidos.

Evaluation. En esta parte, me encargué de evaluar la potencia de entrega y de mejora, así como de definir las acciones a seguir basándonos en las potencias de nuestros resultados. Al ver las potencias de entrega y mejora de cada uno de nuestros resultados, en juntas llegamos al acuerdo de deprecuar la primera versión de nuestro modelo, mejorar la segunda y entregar la tercera y final versión.

Deployment. En esta parte mi papel fue el crear los planes de entrega, monitoreo y mantenimiento. También el monitoreo y medición de beneficios.

Luis Ángel

Business understanding. Participé en las discusiones del equipo para determinar los objetivos de negocio y de minería de datos, además, agregué los riesgos identificados por el equipo en una matriz de riesgos, añadiendo sus planes de contingencia y mitigación. En esta etapa aprendí sobre la importancia de la comunicación con el Socio Formador, así como del establecimiento de los puntos iniciales del proyectos a modo de base.

Data understanding. Me encargué del análisis de técnicas para el manejo de los datos, para ello, realicé un script que separaba las imágenes en clases y en conjuntos de prueba, entrenamiento y validación. También realicé un script que implementa K-fold cross validation como técnica alternativa del uso de datos durante el entrenamiento. Durante esta etapa, aprendí y apliqué técnicas que son de gran

utilidad al analizar datos, además, pude ver más acerca de big data y de computación en la nube.

Data preparation. Participé en la clasificación manual de los datos y en la toma de decisiones para definir cómo se realizaría el proceso de modelado y cómo sería la experiencia de usuario con el proyecto. También me encargué de unir los datos clasificados por cada integrante. En este punto aprendimos sobre el manejo y uso de datos ya de manera aplicada a una situación en la que contamos con muchos datos sin procesar.

Modeling. Brindé un primer modelo que había utilizado a forma de prueba y que fue de utilidad para comenzar con el proceso. Realicé el diseño de pruebas, y contribuí en la definición de las métricas requeridas para poder justificar correctamente los criterios de éxito, así como en el análisis de resultados del último modelo, además, elaboré una introducción general al modelado que permite explicar nuestro proceso seguido en las diferentes iteraciones. Desarrollé, en conjunto con Oskar, la interfaz gráfica de usuario que permite realizar la selección de las camas. Esta fue la etapa que me dejó un mayor aprendizaje, ya que apliqué conocimientos de Machine Learning y de desarrollo de software aplicado a un problema de la vida real.

Evaluation. Participé en la redacción general del documento, con un mayor énfasis en la evaluación de los resultados de minería de datos y en la revisión del proceso. Esto me ayudó a generar conclusiones que utilicen tanto el lenguaje técnico utilizado en la fase de modelado, como un lenguaje más común que pueda justificar al Socio Formador que se lograron los objetivos deseados. Además, hacer una revisión del proceso me fue de utilidad para identificar áreas de mejora.

Deployment. Colaboré en el desarrollo de los ejecutables de la interfaz de usuario, además de participar en la redacción del reporte final y del plan de entrega, documentando nuestra experiencia como equipo y la experiencia del Socio Formador. Esto me permitió conocer de mejor manera todo lo que es necesario entregar en un proyecto de minería de datos, reconociendo la importancia de la documentación correcta, además de que me permitió hacer una reflexión final sobre todo lo que he aprendido a lo largo del proyecto.

Uri Gopar:

Business Understanding: Participe activamente con todos los integrantes del equipo para la definición correcta de nuestros objetivos de negocio, además de la redacción del documento y creación de nuestro plan inicial.

Data Understanding: Fui el responsable de realizar nuestra carga de datos a un servidor de AWS para probar el funcionamiento correcto de nuestro almacenamiento

y flujo de datos, de igual forma realicé un pequeño tutorial de los pasos a seguir para la creación de nuestra arquitectura planteada.

Data Preparation: Me encargué clasificar algunas imágenes y de analizar los registros generados al terminar el proceso de clasificación, redactando en el documento los archivos totales con de nuestros diferentes tipos de clasificaciones (vaca acostada, vaca parada y cama vacía).

Modeling: Me vi involucrado en la fase de pruebas del modelo, donde con Oskar y Toño detectamos errores de predicciones, participé en la creación de mejoras para el modelo 3. Realicé con Toño la prueba arquitectura, la cual descubrimos los requerimientos mínimos de hardware que se necesitarán para el correcto funcionamiento de nuestro modelo en el CAETEC. Con Oskar realizamos el conjunto de análisis de los datos finales los cuales ayudarán a nuestro socio formador para la toma de decisiones.

Evaluation: Participe en la creación de gráficas las cuales fueron de apoyo para medir el rendimiento de nuestros modelos en situaciones reales y de esta manera tener fundamentos para la elección de nuestro modelo, también proporcione la información necesaria para saber el cumplimiento de nuestro objetivo de negocio.

Deployment: Para esta etapa colabore con la realización de Plan de entrega, de igual forma con Mafer nos encargamos de crear el documento de Monitoreo y mantenimiento de nuestra solución y realice la entrevista con el socio formador conocer su satisfacción con este proyecto y el método de entrega de nuestra solución y me encargue de redactar junto con Luis la experiencia que tuvimos como equipo en este semestre.

Mafer Moreno:

Business Understanding: Participé con mi equipo para definir la necesidad del cliente y establecer el objetivo, colaborando con las ideas de la posible solución. También, me encargué de hacer el resumen de negocio de nuestro socio formador y definir el porqué es relevante el trabajo que estamos haciendo. Ayudé a definir los criterios de éxito, establecí con Gama el mapeo de recursos e hice el glosario de terminologías. Revisé ortografía y redacción.

Data Understanding: Describí las herramientas y tecnologías que vamos a utilizar, además de especificar el porqué las vamos a utilizar y también hice el reporte inicial de la colecta de datos. Revisé ortografía y redacción en algunas partes del documento.

Data Preparation: Hice la introducción del documento, clasifiqué 363 imágenes en los tres estados posibles de las camas que son “vaca acostada”, “vaca de pie” y

“cama vacía”. Hice el reporte de la construcción del dataset. Participé en la decisión sobre descartar algunas imágenes y establecer criterios de descarte de las imágenes. Revisé ortografía y redacción del documento.

Modeling: Hice los documentos para las tres etapas de modelado que tuvimos. En los tres documentos de modeling, hice la introducción, técnicas de modelado, diseño y configuración de los modelos, transformaciones de imágenes, entrenamiento, métricas y pruebas y las evaluaciones de los modelos. Redacté resultados del análisis de datos del Modelo 2, también, en los 3 modelos, me encargué del entrenamiento del modelo, evaluando los modelos para que estos cumplieran con nuestros objetivos de minería de datos y con ello, establecer los posibles parámetros que se pudieran mejorar de un modelo para la mejora del siguiente. Hice las tablas comparativas de los modelos.

Evaluation: Hice la introducción, correcciones de redacción y gramática, la tabla comparativa de modelos, describí el proceso de CRISP-DM para modeling y redacté los tiempos de carga y de ejecución de nuestro modelo, con su respectivo reporte del porqué estos tiempos eran buenos en el contexto de nuestro socio formador.

Deployment: Participé en la elaboración del Plan de Entrega detallando resultados así como sus alternativas, además de hacer el Plan de Monitoreo y Mantenimiento con Uri.