

4.º Coloquio en Ciencia de Datos

Proyectos de Incidencia en la Universidad, Industria y Sociedad

11 y 12 de diciembre de 2024

Auditorio «Enrique Valle Flores»



Programa 2025

Jueves 11 de diciembre

6:00 PM	KEVIN GARCÍA MUNDO	Comparación de modelos locales y un modelo global para pronóstico de ventas SKU-tienda.	Este proyecto de trabajo terminal busca mejorar las métricas de evaluación de las ventas en el canal de e-commerce de una cadena de retail mediante pronósticos de ventas semanales a nivel SKU-tienda. Se comparan distintos enfoques de series de tiempo, incluyendo modelos locales clásicos (ARIMA, SARIMA, SARIMAX) y Prophet, frente a un modelo global basado en LightGBM que aprende de forma conjunta de todas las series, así como con los pronósticos actualmente en producción. Como prueba de concepto, se incluye un caso ilustrativo con Análisis Topológico de Datos (TDA) para una sola SKU-tienda. El desempeño se evalúa mediante la reducción del error de pronóstico y la mejora en la disponibilidad de producto, impactando directamente en la experiencia del cliente.
6:20 PM	FELICIANO ROBLES ANDRADE	Detección de patrones de comportamiento de unidades de transporte en la ciudad de Hermosillo, Sonora	El transporte urbano enfrenta problemas que afectan su eficiencia. Este proyecto planea aplicar modelos de inteligencia artificial para predecir tiempos de llegada y detectar anomalías. Su innovación está en la creación de nuevos parámetros de datos, captando el contexto informal del transporte. El objetivo es lograr una predicción más precisa que los métodos simples, utilizando una robusta y masiva base de datos disponible.
6:40 PM	ANGEL ALBERTO BARRAZA REAL	Predicción de tiempo de cruce por vehículo con drones y YOLO mediante aprendizaje continuo vuelo a vuelo en el borde	Se presentará un sistema de análisis de tráfico urbano usando videos de dron y visión por computador. El sistema detecta y sigue vehículos con YOLO y tracking multiobjeto, para calcular cuánto tiempo tarda cada auto en cruzar un semáforo. Con esos datos entreno un modelo ligero que predice el tiempo de cruce restante por vehículo, actualizándolo vuelo a vuelo con estrategias sencillas de aprendizaje continuo en un escenario tipo Edge.
Café			
7:20 PM	FEDERICO MIGUEL CIRETT TORRES	Aplicación de PINNs para ajuste óptimo de curvas de crecimiento en granos de kéfir tratados con ultrasonido	El trabajo de investigación tiene como objetivo principal determinar el ajuste óptimo de curvas de crecimiento correspondientes a granos de kéfir sometidos a tratamiento por ultrasonido. Para abordar esto, se propone la implementación de Physics-Informed Neural Networks (PINNs), las cuales incorporan conocimiento físico dentro del proceso de aprendizaje de una red neuronal. Este enfoque permite modelar de manera robusta el comportamiento dinámico del crecimiento.
7:40 PM	JESÚS MANUEL SOLÍS DURÁN	Explicabilidad como marco de entendimiento y ajuste de modelos para pronóstico de series de tiempo	En esta charla nos centraremos en el uso de la explicabilidad de modelos de aprendizaje profundo para pronóstico de series de tiempo, para ajustar criterios e hiperparámetros no-triviales en el proceso de entrenamiento, a fin de aportar confiabilidad al ciclo de pronóstico.
8:00 PM	JULIO WAISSMAN VILANOVA	Presentación de posibles proyectos de trabajo terminal para alumnos de primer semestre	
Canapés			

Viernes 12 de diciembre

6:00 PM	MANUEL COTA SANTAEZ	Modelo y visualización de brote de dengue en Hermosillo	Con este proyecto se calibrará un modelo basado en ecuaciones diferenciales ordinarias para describir la transmisión de la fiebre del dengue, tanto en su forma clásica como hemorrágica. Los datos utilizados incluyen una componente de georreferenciación indexada por fechas de confirmación. Con esta información, se busca realizar una simulación georreferenciada basada en agentes que permita describir la progresión del dengue en al menos dos clústeres dentro de la ciudad de Hermosillo.
6:20 PM	ARIAN MILANÉS GARCÍA	Desarrollo y comparación de modelos para la predicción de demanda de energía a corto plazo en zonas de carga de la Gerencia del Control de Noroeste	El objetivo es la comparación de modelos de predicción de demanda de energía a corto plazo en Zonas de Carga de la Gerencia Regional Noroeste del CENACE, utilizando la información histórica de demanda de energía, variables meteorológicas y de calendario. El proceso consistirá en primer lugar, en replicar un proceso ya desarrollado para la estimación de la demanda a nivel Gerencia de Control con un modelo tipo seq2seq, adaptándolo a la ingesta y procesamiento de datos a nivel Zona de Carga, para posteriormente implementar diferentes modelos basados en mecanismos de atención.
6:40 PM	GERARDO LEYVA CONDE	Infraestructura para automatización de reportes de eficiencia de alimentación de camarones.	Creación de un pipeline para la obtención de información desde diversas fuentes de datos, con el fin de integrarlas en un solo repositorio y, posteriormente, visualizarlas mediante un dashboard interactivo.
Café			
7:20 PM	EDDEL ELI OJEDA AVILES	Sistema de visión por computadora para la detección y clasificación de productos perecederos en la industria de retail.	El objetivo del proyecto es desarrollar un sistema de visión por computadora capaz de detectar y clasificar frutas y verduras según su estado de maduración y posibles defectos presentes, por medio de fotos y modelos basados en la arquitectura Detectron2 con segmentación de instancias. Estos modelos serán entrenados con imágenes correctamente etiquetadas mediante la herramienta CVAT. Esta solución permitirá una inspección continua y precisa en el área de recibo de cada sucursal, mejorando la eficiencia operativa actual a través de la generación automática de reportes con las métricas señaladas en la ficha técnica de cada producto, junto con la recomendación de aceptar o rechazar el pedido. El principal beneficio será la reducción de la merma por productos deteriorados originadas por error humano en la inspección del producto en el área de recibo.
7:40 PM	CATHERINE LEE SCOTT LUEVANO	Segmentación de mercado y optimización de ubicaciones para nuevos Drives	CAFFENIO enfrenta el reto de seguir expandiendo su red de Drives en un entorno competitivo donde las decisiones de apertura se basan principalmente en la experiencia y el juicio humano, a pesar de contar con información interna de ventas, lealtad y ubicación geográfica que no se aprovecha de manera integrada. El objetivo de este trabajo terminal es diseñar y evaluar un producto de datos que permita segmentar clientes y clasificar puntos de venta, así como apoyar la selección de nuevas ubicaciones con base en evidencia. Para ello se integrarán bases internas de la empresa (ventas históricas, programa de lealtad MiCaffenio, catálogo de Drives) con información demográfica y geográfica pública, como datos de INEGI y servicios de mapas digitales. Se utilizarán técnicas de análisis exploratorio, segmentación mediante métodos de clustering y análisis multicriterio de ubicación, implementados principalmente con Python y SQL/BigQuery, y los resultados se plasmarán en tableros interactivos en Power BI dirigidos a las áreas de Expansión, Mercadotecnia, Lealtad y Operaciones. Se espera que este proyecto contribuya a decisiones de expansión más precisas, rentables y rápidas, reduciendo el riesgo de inversiones no óptimas y fortaleciendo la cultura de decisiones basadas en datos dentro de CAFFENIO
8:00 PM	CELIA HERNÁNDEZ CRUZ	Modelo analítico para la optimización del pedido sugerido en puntos de venta del sector alimentos y bebidas	El proyecto desarrolla un modelo analítico-operativo para optimizar el pedido sugerido en los puntos de venta de CAFFENIO, integrando datos de SAP, POS y fuentes externas como clima, promociones y eventos. Hoy, el proceso de inventarios presenta inconsistencias, ausencia de trazabilidad y dependencia de herramientas manuales, lo que provoca sobreinventario, mermas y quiebres de stock. La propuesta construye un pipeline analítico confiable y reproducible en BigQuery, donde se procesan ventas, inventarios, caducidades y variables externas. El modelo predictivo se desarrollará en Python, y los resultados se visualizarán mediante Power BI para la validación operativa y toma de decisiones.
Convivio y cierre			



"El saber de mis hijos hará mi grandeza"



MCD

Maestría en Ciencia de Datos