

Reporte # 2 Microproyecto

Pronóstico de ventas para una máquina expendedora de café
Curso: Proyecto - Desarrollo de Soluciones
Grupo # 4

1. Resumen del problema

1.1. Contexto

El proyecto se centra en el análisis y pronóstico de ventas de una máquina expendedora de café ubicada en un centro comercial de Vinnytsia, Ucrania. Se trata de un entorno de alta concurrencia, donde la gestión eficiente del inventario y la anticipación de la demanda resultan críticas para mantener la rentabilidad y garantizar la disponibilidad de productos. El espacio de almacenamiento en la máquina es limitado y la reposición inmediata no siempre es posible, por lo que el manejo adecuado de la demanda es un factor clave.

1.2. Problema identificado

En el diagnóstico inicial se reconocieron tres dificultades principales que limitan la gestión de la máquina:

- **Falta de visión consolidada de ventas:** no se cuenta con un sistema que identifique claramente qué productos, franjas horarias o métodos de pago generan la mayor proporción de ingresos.
- **Subutilización de información temporal:** los registros por hora, día de la semana y método de pago no se aprovechan para optimizar inventario, planear mantenimientos o diseñar promociones específicas.

- **Ausencia de mecanismos de pronóstico:** la reposición de inventario se realiza de manera reactiva, lo que provoca desabastecimientos en horas pico, exceso de stock en productos de baja rotación y disminución en la satisfacción del cliente.

1.3. Pregunta de negocio

La pregunta que guía el proyecto es: **¿cuáles son los patrones de demanda por producto, hora y método de pago que permitan pronosticar ventas a corto plazo (1-7 días) y optimizar la reposición de la máquina expendedora?**

De esta se derivan cuestiones adicionales, como: qué productos concentran la mayor rotación en distintas franjas horarias, cómo se diferencia el comportamiento de compra según el método de pago y qué velocidad de venta minimiza el riesgo de quiebres de inventario.

El por qué se limitó el horizonte a 7 días

Robustez del modelo: Con solo un año de datos históricos, horizontes mayores incrementarían exponencialmente la incertidumbre. La degradación predictiva en forecasting de series temporales es no-lineal después de 7 días.

Contexto de negocio: Para una máquina expendedora, 7 días es suficiente para ciclos de reposición y captura patrones semanales recurrentes. Sin embargo, para la última entrega de este proyecto se analizará nuevamente la decisión.

1.4. Alcance

El proyecto se limita a generar pronósticos de ventas diarios con un horizonte de hasta siete días. El análisis se orienta a productos individuales y totales agregados, y contempla la validación de distintos modelos predictivos utilizando métricas de error como RMSE, MAE y MAPE. El resultado esperado es un tablero interactivo que permita consultar tanto los patrones históricos como las predicciones, incorporando además estimaciones de inventario y costos operativos.

Quedan fuera del alcance los pronósticos de mediano y largo plazo, la integración con sistemas externos de punto de venta, la fijación de precios dinámicos y el análisis de competencia. Asimismo, se reconoce la limitación de disponer únicamente de un año de observaciones.

1.5. Datos a emplear

La base de datos principal proviene del archivo `index_1.csv` (Kaggle), con 3.636 transacciones

registradas entre marzo de 2024 y marzo de 2025. El dataset incluye variables como: fecha de transacción, hora exacta, método de pago (tarjeta o efectivo), monto de la compra y nombre de la bebida adquirida. De manera complementaria se incorporan calendarios oficiales de festivos en Ucrania y, de ser posible, datos meteorológicos de Vinnytsia en el mismo período, con el fin de capturar variaciones externas que puedan afectar la demanda.

2. Modelos desarrollados y su evaluación

Con el propósito de construir un sistema de pronóstico de la demanda de café, se implementó un conjunto diverso de modelos que abarcan enfoques clásicos, estadísticos y de aprendizaje profundo. La meta consistió en capturar las dinámicas temporales de consumo a un nivel de granularidad diaria y por producto, permitiendo generar proyecciones útiles para la planificación de inventarios y compras.

El proceso de evaluación se realizó mediante un esquema de *backtesting* con orígenes rodantes. Esta metodología fue seleccionada por su capacidad de reproducir condiciones similares a las de un despliegue real, reduciendo el riesgo de fuga de información y garantizando que las métricas de error reflejaran el desempeño esperado en producción. En cada iteración, el modelo se entrenó con una ventana fija de datos históricos y se evaluó sobre un horizonte de siete días, repitiendo el procedimiento a lo largo de todo el periodo disponible.

La variable objetivo se definió como la cantidad diaria de unidades vendidas, en lugar de los ingresos monetarios. De esta manera, se aisló el comportamiento de la demanda respecto a fluctuaciones de precios y se evitó introducir ruido exógeno en el entrenamiento. El cálculo de ingresos se incorporó únicamente en etapas posteriores, multiplicando las unidades proyectadas por el precio promedio observado en cada bebida.

Se incorporaron variables de calendario (día de la semana, mes, feriados) y transformaciones rezagadas de la serie, asegurando que en ningún caso se utilizara información futura. Para las variables climáticas se consideró su disponibilidad potencial en tiempo real, manteniendo consistencia con un escenario operativo. La construcción de variables rezagadas y ventanas móviles se diseñó cuidadosamente para no comprometer la separación entre entrenamiento y prueba.

Los modelos probados incluyeron tres líneas base: el ingenuo (último valor observado), el estacional semanal (mismo día de la semana anterior) y un promedio móvil de siete días. Estos sirvieron como punto de referencia mínimo para determinar la validez de los métodos más avanzados. Entre los modelos principales se incluyeron: *LightGBM* en un esquema directo multi-horizonte, *SARIMAX* con regresores exógenos, *Prophet* con componentes de tendencia y estacionalidad, y una red neuronal *LSTM* configurada para emitir múltiples horizontes de salida simultáneamente. Cada

modelo fue entrenado bajo criterios homogéneos de validación y se ajustaron sus hiperparámetros para optimizar el desempeño en el horizonte de siete días.

El entrenamiento de la LSTM se apoyó en técnicas de regularización como *EarlyStopping*, con el fin de prevenir sobreajuste. En el caso de los modelos basados en gradiente, se aplicaron estrategias de reducción de dimensionalidad para priorizar variables con mayor relevancia y evitar redundancias.

El *SARIMAX* se estimó con parámetros ajustados para capturar estacionalidad semanal y efectos de calendario. Prophet, por su parte, permitió una calibración rápida de estacionalidad y tendencia, aunque con limitaciones frente a dinámicas complejas.

Las métricas seleccionadas fueron el error absoluto medio (MAE), la raíz del error cuadrático medio (RMSE) y el porcentaje absoluto medio simétrico (sMAPE). Se buscó un balance entre precisión absoluta y relativa, de manera que la selección del modelo final priorizara un bajo sMAPE sin comprometer los valores de MAE y RMSE.

3. Observaciones y conclusiones sobre los modelos

El análisis de resultados mostró que las líneas base confirmaron la presencia de un patrón semanal en la demanda, aunque sus niveles de error fueron demasiado elevados para su uso práctico. Su utilidad radicó en servir como piso de comparación, que posteriormente permitió dimensionar las ganancias de los modelos más avanzados. El modelo en conjunto (ensemble), construido a partir de la combinación ponderada de predicciones, ofreció un desempeño superior con un sMAPE cercano a 72, además de errores absolutos y cuadráticos menores que los de cualquier modelo individual. Este resultado confirmó que la agregación de enfoques heterogéneos reduce errores extremos y ofrece mayor estabilidad, evitando depender de un único método.

En cuanto a los modelos individuales, el *SARIMAX* se destacó por su interpretabilidad y competitividad en MAE y RMSE, aunque presentó dificultades en periodos de baja demanda donde la precisión relativa cayó. Prophet ofreció facilidad de uso y rapidez, pero no alcanzó la exactitud requerida para el nivel de granularidad del análisis. La red *LSTM* logró capturar dependencias temporales y presentó buenos resultados en errores absolutos, pero el tamaño limitado de la muestra redujo su capacidad de generalización y afectó su desempeño relativo. Por su parte, *LightGBM* mostró un rendimiento más débil de lo esperado en sMAPE, lo que sugiere la necesidad de ajustar su especificación, revisar las variables utilizadas y explorar funciones de pérdida adaptadas a series temporales.

La evidencia respalda al modelo ensemble como la alternativa más sólida para un eventual despliegue operativo, ya que combina la robustez de modelos paramétricos, la sensibilidad de redes

neuronales y la flexibilidad de enfoques de aprendizaje automático. Los modelos de referencia se conservarán como herramientas de control y comparación, mientras que SARIMAX y LSTM aportan componentes complementarios de interpretabilidad y captura de patrones complejos. Prophet se mantiene como herramienta útil para exploración y visualización, y LightGBM requiere una re-especificación para mejorar su competitividad en este contexto.

4. Gestión de experimentos con MLflow

Para garantizar la trazabilidad, organización y comparación sistemática de los modelos, se implementó un flujo de trabajo con *MLflow*. Esta herramienta permitió registrar de manera automática cada experimento, almacenando tanto los parámetros utilizados en el entrenamiento como las métricas obtenidas en la validación y los artefactos resultantes.

El diseño del registro consideró una estructura jerárquica en la que cada modelo entrenado se asociaba a una ejecución (*run*) identificada por un conjunto único de metadatos. En cada ejecución se almacenaron hiperparámetros relevantes, como el horizonte de predicción, la cantidad de rezagos, las funciones de pérdida empleadas o la configuración de regularización. De forma paralela, se registraron las métricas de desempeño, lo que permitió establecer comparaciones objetivas entre modelos y reproducir resultados en el futuro.

Los artefactos, que incluyeron gráficos de validación, series de pronóstico y el propio modelo serializado, se guardaron automáticamente en el directorio correspondiente a cada experimento. Este procedimiento facilitó el acceso posterior a resultados específicos y eliminó la dependencia de procesos manuales para el almacenamiento de archivos. Gracias a esta sistematización, fue posible documentar de manera transparente el ciclo completo de entrenamiento y evaluación, garantizando que cualquier modelo pudiera ser recuperado y auditado.

Un aspecto adicional de la integración con MLflow fue la definición de un registro central de modelos. Una vez seleccionada la mejor especificación en función de las métricas, el modelo se promovió a un estado estable dentro del registro, desde donde quedó disponible para ser cargado en el tablero de pronósticos. Esta práctica permitió mantener un control estricto sobre la versión del modelo en uso y ofreció un mecanismo sencillo para actualizarlo cuando se identificaran mejoras en entrenamientos posteriores.

En la Figura 1 y 2 se muestra la interfaz de MLflow utilizada en el proyecto, donde se registraron automáticamente las ejecuciones con sus parámetros y métricas correspondientes.

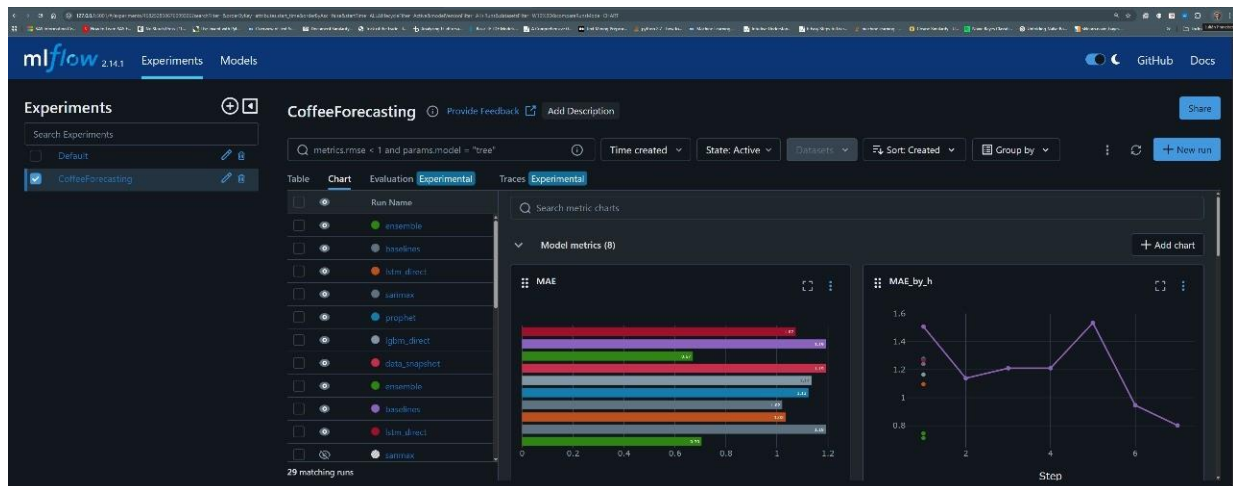


Figura 1: Evidencia del uso de MLFlow.

Fuente: Elaboración propia.

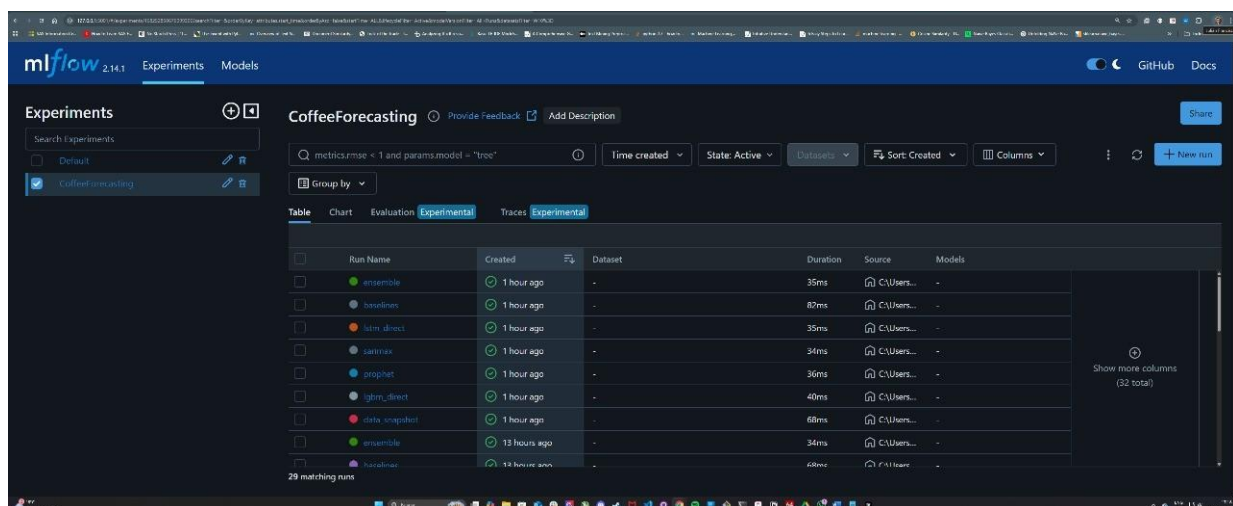


Figura 2: Evidencia del uso de MLFlow.

Evidencia del uso de MLFlow.

5. Tablero desarrollado

El tablero interactivo fue diseñado para dar respuesta a la pregunta de negocio planteada en el proyecto: **¿cuáles son los patrones de demanda por producto, hora y método de pago que permitan pronosticar ventas a corto plazo y optimizar la reposición de la máquina expendedora?**

5.1. Cumplimiento del alcance

El tablero se enfoca en pronósticos de corto plazo con horizonte de hasta siete días, utilizando como insumo los datos históricos de la máquina expendedora entre marzo de 2024 y marzo de 2025. La granularidad es diaria y por producto, y se incluyen métricas de desempeño (RMSE, MAE, MAPE) que permiten evaluar la calidad de las predicciones. De este modo, el desarrollo se ajusta al alcance definido: predicciones de corto plazo, sin extenderse a proyecciones de mediano o largo plazo ni a análisis externos al contexto de la máquina.

5.2. Predicciones generadas

El tablero ofrece estimaciones tanto del total de ventas como del detalle por producto. Los resultados se visualizan de manera comparativa, incluyendo validaciones contra el comportamiento histórico. Además, se derivan estimaciones de insumos y costos asociados, lo que permite anticipar requerimientos de inventario y proyectar ingresos y utilidades.

5.3. Interfaz adecuada para la toma de decisiones

La interfaz se diseñó bajo criterios de claridad y usabilidad. Los filtros en la barra lateral permiten seleccionar fácilmente bebida, rango de fechas, día de la semana, franja horaria y modelo de pronóstico. El contenido central se organiza en módulos que muestran: indicadores clave de desempeño (ventas recientes, variaciones porcentuales, producto más vendido), análisis de patrones de compra, resultados de pronósticos, estimaciones de inventario y evaluación financiera. El estilo visual uniforme, con colores neutros y tarjetas compactas, favorece la lectura y la interpretación rápida de la información.

5.4. Aporte a la pregunta de negocio

Con la integración de estos elementos, el tablero demuestra que es posible identificar patrones de consumo diferenciados por producto, horario y día de la semana, y generar predicciones confiables para un horizonte de hasta siete días. La información obtenida apoya la planificación de la reposición de la máquina expendedora, reduce riesgos de desabastecimiento o sobre stock y permite evaluar el impacto financiero de las decisiones. En conjunto, el tablero cumple la pregunta de negocio planteada y materializa el alcance del proyecto en una herramienta práctica de apoyo a la gestión.

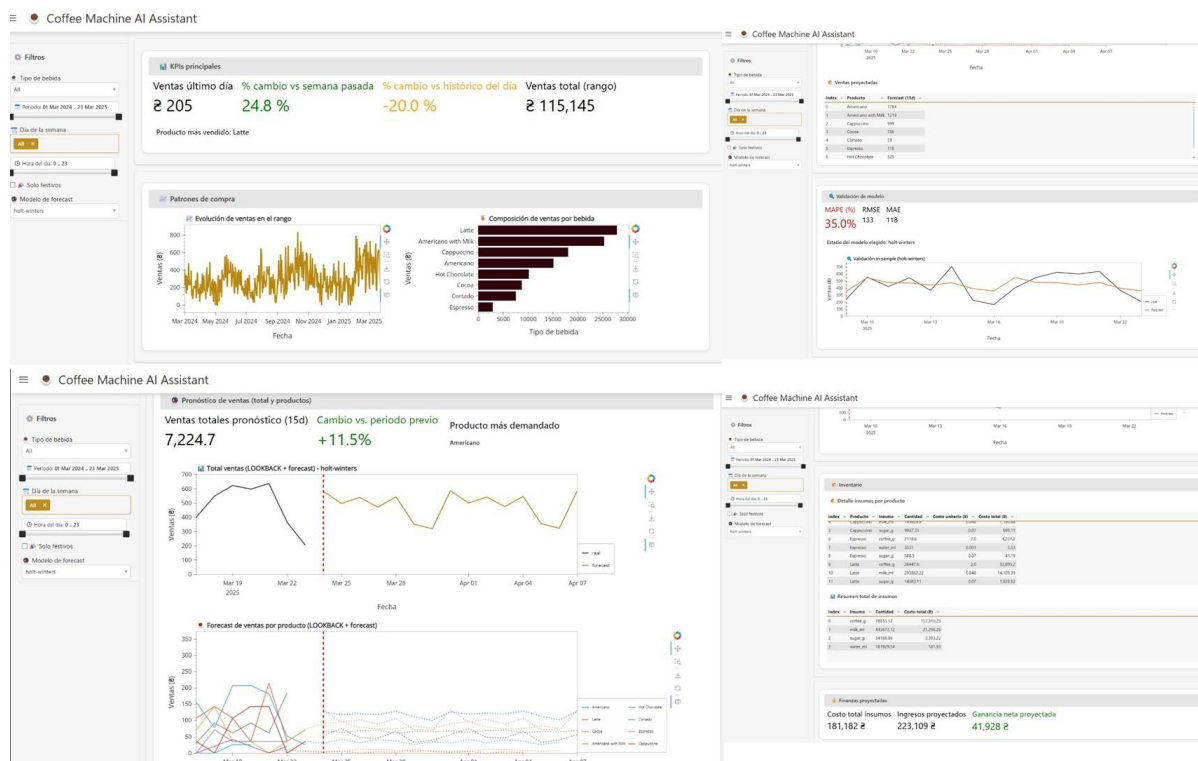


Figura 3: Dashboard desarrollado y sus distintas interfaces y capacidades.

Fuente: Elaboración propia.

6. Reporte de trabajo en equipo

El desarrollo del proyecto se llevó a cabo de manera colaborativa mediante el uso de control de versiones con Git y la plataforma GitHub. Cada integrante trabajó en ramas independientes, lo que permitió organizar los avances y reducir conflictos al momento de integrar los cambios. En particular, todos los miembros exploraron de manera individual la integración de *MLflow* dentro de sus ramas de trabajo, pero la consolidación final se realizó en el tablero de pronósticos con una sola implementación unificada.

La distribución de las actividades fue la siguiente:

- **Jaime Vera:** construcción e implementación del *dashboard*, integración de métricas y conexión final con el modelo ganador en MLflow, junto con las especificaciones correspondientes.

Elaboración del reporte técnico.

- **Julián Cujabante:** Diseño del *pipeline* principal de entrenamiento, ajuste de especificaciones, control de versiones de modelos y descripción detallada del modelo final.
- **Uldy Paloma:** Desarrollo y versionamiento de modelos, realización de pruebas adicionales con métricas de negocio.
Investigación y elaboración de variables exógenas relacionadas con festivales y datos meteorológicos de Vinnytsia. Creación diccionario de datos correspondiente y ajustes al reporte.
- **Rafael Ortega:** Pruebas adicionales de entrenamiento y validación.

La integración de los aportes se realizó mediante *pull requests*, asegurando que cada contribución fuera validada antes de incorporarse al repositorio compartido. El esquema de ramas permitió mantener un flujo de trabajo transparente y trazable, con responsabilidades definidas en cada etapa.

Enlace repositorio: <https://github.com/maia-proyecto-integrador/coffee-sales-project.git>

La Figura 4 presenta un ejemplo del historial de commits y ramas, donde se evidencia la participación activa de los integrantes y la organización de las contribuciones en ciclos de desarrollo.

