



Aplicación de inteligencia artificial a la planificación financiera de productos secos en el ciclo III: demanda, márgenes y simulación de escenarios

Autor:

Mg. Emiliano Daniel Iparraguirre

Director:

Título y Nombre del director (pertenencia)

*Esta planificación fue realizada en el curso de Gestión de proyectos
entre el 29 de abril de 2025 y el 17 de junio de 2025.*

Índice

1. Descripción técnica-conceptual del proyecto a realizar	5
2. Identificación y análisis de los interesados	6
3. Propósito del proyecto	7
4. Alcance del proyecto	7
5. Supuestos del proyecto.	8
6. Product Backlog	9
7. Criterios de aceptación de historias de usuario	11
8. Fases de CRISP-DM	16
9. Desglose del trabajo en tareas	17
10. Diagrama de Gantt	24
11. Planificación de Sprints	29
12. Normativa y cumplimiento de datos (gobernanza)	37
13. Gestión de riesgos	39
14. Sprint Review	43
15. Sprint Retrospective	43

Registros de cambios

Revisión	Detalles de los cambios realizados	Fecha
0	Creación del documento	29 de abril de 2025
1	Se completa hasta el punto 5 inclusive	13 de mayo de 2025
2	Se completa hasta el punto 9 inclusive	20 de mayo de 2025
3	Se completa hasta el punto 12 inclusive	27 de mayo de 2025
4	Se completa el plan	3 de junio de 2025
5	Se realizan las correcciones finales	16 de junio de 2025

Acta de constitución del proyecto

Buenos Aires, 29 de abril de 2025

Por medio de la presente se acuerda con el Mg. Emiliano Daniel Iparraguirre que su Trabajo Final de la Carrera de Especialización en Inteligencia Artificial se titulará “Aplicación de inteligencia artificial a la planificación financiera de productos secos en el ciclo III: demanda, márgenes y simulación de escenarios” y consistirá en desarrollar un modelo basado en inteligencia artificial para predecir la demanda y los márgenes brutos de productos secos en Cagnoli S.A., integrándolo al proceso de planificación financiera para mejorar la toma de decisiones y simular escenarios. El trabajo tendrá un presupuesto preliminar estimado de 677 horas y un costo de USD 27.080, con fecha de inicio el 29 de abril de 2025 y fecha de presentación pública el **xx de xxx de 2026**.

Se adjunta a esta acta la planificación inicial.

Dr. Ing. Ariel Lutenberg
Director posgrado FIUBA

Santiago Pagola
Cagnoli S.A.

Título y Nombre del director
Director del Trabajo Final

1. Descripción técnica-conceptual del proyecto a realizar

El presente proyecto se realiza en el marco de la Especialización en Inteligencia Artificial, y surge a partir de la experiencia profesional del alumno como consultor en planificación financiera e inteligencia de negocios en Cagnoli S.A., una empresa argentina con más de un siglo de trayectoria en la industria de alimentos. La firma se especializa en la elaboración de fiambres y embutidos con estándares de calidad certificados, y cuenta con distintas unidades productivas organizadas en ciclos. En particular, este trabajo se enfoca en el **Ciclo III** (producción de fiambres y chacinados), correspondiente a la producción de productos secos (salamines), una de las líneas más relevantes por su imagen de marca, volumen de ventas y contribución al margen.

El objetivo del proyecto es diseñar una herramienta de planificación financiera basada en Inteligencia Artificial, orientada a pronosticar la demanda y estimar márgenes brutos por producto y canal. Además, se desarrollará un módulo de simulación de escenarios operativos y financieros que permita evaluar el impacto ante posibles variaciones en factores críticos como mermas productivas, precios de insumos, regulaciones o cambios en el patrón de consumo. Esta herramienta busca integrarse al proceso de planificación mensual que coordina las áreas Comercial, Producción y Finanzas, con el fin de mejorar así la calidad de las decisiones.

No existen actualmente acuerdos de confidencialidad o financiamiento externo. Los entregables serán propiedad del alumno, y se resguardará la confidencialidad comercial en caso de futuras publicaciones académicas.

Contexto operativo

El proceso de elaboración de productos secos implica tiempos prolongados de maduración y secado (que pueden superar los 30 días), por lo que una planificación incorrecta genera consecuencias graves: desabastecimiento, vencimientos, capacidad ociosa o pérdida de clientes. Además, se enfrenta una alta estacionalidad: la demanda se concentra en verano, con marcada caída en los meses fríos.

La empresa, si bien ha incorporado recientemente herramientas de control de gestión y planificación, aún no cuenta con modelos predictivos robustos que integren física y financieramente el negocio.

Aspectos técnicos y estado del arte

La solución se encuentra en una etapa temprana del ciclo de vida: diseño conceptual, análisis del problema y planificación del desarrollo. Durante el proyecto se avanzará hasta el desarrollo funcional del modelo, su evaluación en entorno de pruebas y su validación con datos reales, sin incluir en el alcance aún la integración de la solución con los sistemas transaccionales de la empresa.

En la figura 5 se presenta el diagrama en bloques del sistema propuesto. Se observa cómo, a partir de diversas fuentes de datos (ventas históricas, costos variables, eventos comerciales y variables externas), el sistema procesa la información mediante tres módulos: uno de predicción de demanda (basado en SARIMA, Prophet y algoritmos de *machine learning* supervisado como XGBoost y Random Forest), uno de estimación de márgenes (que calcula costos y contribución marginal proyectada), y otro de simulación de escenarios (utilizando árboles de decisión y

análisis "what if"). La salida de este procesamiento se visualiza mediante Power BI o Streamlit, generando reportes que sirven como insumo clave para las reuniones mensuales de planificación operativa y financiera.

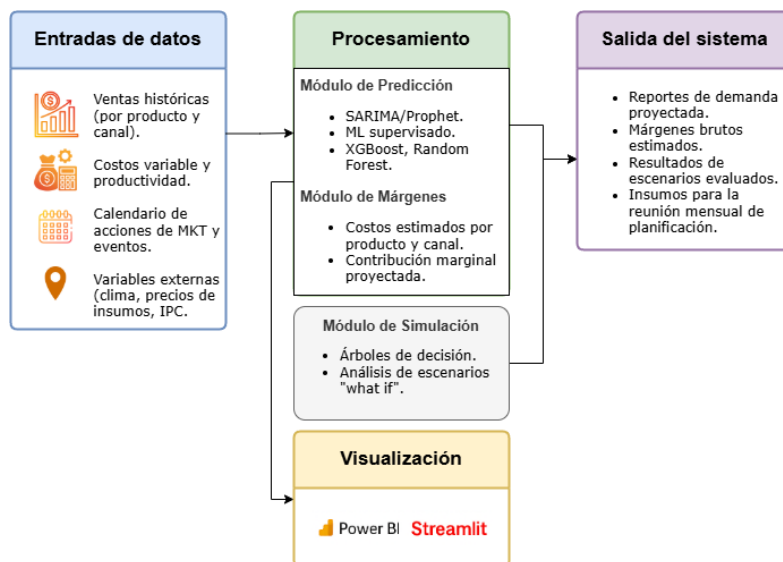


Figura 1. Diagrama en bloques del sistema.

Cliente y valor esperado

El principal cliente del proyecto es el equipo de planificación, ya que requiere información confiable y anticipada para mejorar sus decisiones. La innovación del proyecto radica en conectar el planeamiento operativo y financiero con modelos predictivos y simulación en una plataforma visual y accesible. La solución permitirá analizar semanalmente las proyecciones de demanda y márgenes por canal, junto con simulaciones de impacto financiero ante cambios en precios de insumos o contexto de mercado. Será utilizada como insumo objetivo en las reuniones mensuales de planificación comercial y presupuestaria.

2. Identificación y análisis de los interesados

Rol	Nombre y Apellido	Organización	Puesto
Cliente	Santiago Pagola	Cagnoli S.A.	Responsable de Planificación
Impulsor	Juan Pedro Cagnoli	Cagnoli S.A.	Gerente de Operaciones
Responsable	Emiliano Iparraguirre	FIUBA	Alumno
Orientador	Título y Nombre del director	pertenencia	Director del Trabajo Final
Usuario final	Analistas de Planificación	Cagnoli S.A.	-

A continuación, se listan las principales características de cada interesado:

- Cliente: el Cr. Santiago Pagola colaborará con la explicación del proceso de planificación, la definición de los requerimientos y la coordinación de ejecución de pruebas de aceptación.
- Impulsor: el Ing. Juan Pedro Cagnoli es el principal referente y sponsor de este proyecto dentro de la organización. Esta acción se inserta en un conjunto de iniciativas estratégicas que él mismo ha impulsado, con el propósito de perfeccionar la planificación, aumentar la precisión de los pronósticos y sustentar las decisiones empresariales en modelos objetivos.
- Orientador: **características del orientador.**
- Usuario final: los analistas del Departamento de Planificación son los usuarios finales de esta herramienta, antes y durante las reuniones de planificación, junto al resto de las áreas involucradas (Comercial, Producción, Logística y Compras).

3. Propósito del proyecto

El propósito de este proyecto es desarrollar un modelo predictivo basado en inteligencia artificial que prevea la demanda y los márgenes brutos por producto y canal, integrando datos del sistema ERP de la compañía, costos dinámicos e información comercial. Adicionalmente, como último paso, se incluirá una capa de simulación de escenarios para evaluar impactos de cambios en variables críticas sobre los resultados financieros. Esta solución permite aportar herramientas objetivas para la planificación operativa y financiera integrada en un flujo de trabajo inteligente, proporcionando información confiable y anticipada para la toma de decisiones.

4. Alcance del proyecto

El proyecto se centrará exclusivamente en la familia de productos secos (salamines), por su peso estratégico dentro del portafolio de Cagnoli S.A. Estos productos requieren una planificación diferenciada debido a sus procesos largos de secado y maduración, y a la alta estacionalidad en su demanda.

El presente proyecto incluye:

- El diseño y validación de modelos de predicción de demanda, basados en algoritmos de series temporales (SARIMA, Prophet) y métodos de aprendizaje automático supervisado (XGBoost, Random Forest).
- La estimación proyectada de márgenes brutos por canal, a partir de precios previstos y costos dinámicos asociados a producción y distribución.
- La simulación de escenarios mediante árboles de decisión y análisis de sensibilidad, para evaluar impactos financieros ante variaciones en precios, demanda o restricciones operativas.
- Una prueba de concepto visual (PoC), implementada en Power BI y/o Streamlit, destinada a facilitar la interpretación y evaluación de los resultados por parte de los equipos internos.

El presente proyecto no incluye:

- El entrenamiento formal a usuarios finales o su incorporación en la rutina operativa de la empresa.
- La integración técnica con sistemas transaccionales o con el sistema ERP de la compañía (Microsoft Dynamics).
- La incorporación de otras líneas de productos como cocidos o frescos al análisis, que se encuentran más allá de la familia de productos secos.

5. Supuestos del proyecto

Para el desarrollo del presente proyecto se supone que:

- Se contará con acceso oportuno y completo a los *datasets* históricos necesarios para el modelado: ventas, costos, inventarios, estacionalidad, calendario comercial, entre otros. Los datos estarán en formato utilizable y reflejarán fielmente la realidad operativa del negocio.
- Se prevé que tanto el alumno como los referentes internos de la empresa (Departamento de Planificación) podrán destinar el tiempo necesario para reuniones, revisiones de avances, validación de entregables y provisión de información contextual.
- Se considera que las herramientas previstas (Python, bibliotecas de IA, Streamlit, Power BI, etc.) son adecuadas para el desarrollo del sistema, y que no surgirán restricciones técnicas críticas que impidan ejecutar los modelos o visualizar los resultados.
- El desarrollo se llevará a cabo con los recursos disponibles del alumno (hardware personal, conectividad y acceso a herramientas de código abierto). No se prevé la necesidad de inversiones adicionales en infraestructura o licencias.
- Se asume que no existirán bloqueos internos, resistencias de parte de otras áreas o barreras políticas que interfieran con el acceso a la información o con la validación de resultados por parte del equipo de planificación.
- Para las simulaciones y la interpretación de resultados, se asume que no ocurrirán cambios abruptos en el contexto macroeconómico (inflación, precios regulados, normativas sanitarias) que invaliden los patrones históricos o generen disrupciones graves en la operatoria de la empresa.
- Se considera que los datos históricos reflejan adecuadamente la variabilidad esperada del negocio, y que son válidos para el entrenamiento y evaluación de modelos predictivos.

6. Product Backlog

Los criterios para calcular los story points estarán divididos de la siguiente manera:

- **Dificultad (1–5):** evalúa nivel de esfuerzo técnico y de desarrollo requerido.
- **Complejidad (1–5):** referido al grado de conocimiento y experiencia necesarios.
- **Riesgo (1–5):** pondera la incertidumbre técnica, disponibilidad de datos, o dependencias externas.

Para la ponderación en cada uno de ellos se utilizará la serie de Fibonacci desde el 1 al 13. La escala queda de la siguiente forma: muy bajo (1), bajo (3), medio (5), alto (8) y muy alto (13).

6.1. Épica 1: modelado de demanda

- **Historia de usuario 1:** como analista de planificación operativa, quiero contar con un modelo de predicción de demanda semanal por producto, para anticipar las necesidades de producción.
 - **Dificultad:** medio (5). Requiere diseño, entrenamiento y validación de modelos.
 - **Complejidad:** medio (5). Involucra series temporales con estacionalidad marcada y ruido.
 - **Riesgo:** bajo (3). Depende de la calidad y granularidad de los datos históricos.
 - **Story Points** = $5 + 5 + 3 = 13$.
 - **Prioridad:** alta.
- **Historia de usuario 2:** como responsable comercial, quiero contar con alertas ante desvíos significativos entre demanda proyectada y real para poder ajustar rápidamente la estrategia de ventas.
 - **Dificultad:** bajo (3). Requiere definir umbrales y condiciones de alerta.
 - **Complejidad:** bajo (3). Necesita cálculos comparativos en tiempo real.
 - **Riesgo:** muy bajo (1). El riesgo se acota al mal ajuste de los parámetros de alerta.
 - **Story Points** = $3 + 3 + 1 = 8$ (valor próximo en la serie de Fibonacci).
 - **Prioridad:** media.
- **Historia de usuario 3 (opcional):** como jefe de producción, quiero acceder a predicciones semanales agregadas por línea de producto para facilitar la planificación de turnos y uso de capacidad.
 - **Dificultad:** bajo (3). Se trata de una agregación de las predicciones ya existentes.
 - **Complejidad:** bajo (3). Es una funcionalidad de fácil desarrollo.
 - **Riesgo:** muy bajo (1). Baja dependencia de factores externos.
 - **Story Points** = $3 + 3 + 1 = 8$ (valor próximo en la serie de Fibonacci).
 - **Prioridad:** baja.

6.2. Épica 2: cálculo de márgenes brutos

- **Historia de usuario 4:** como responsable de control de gestión, quiero estimar los márgenes brutos por canal y producto para evaluar la rentabilidad futura.
 - **Dificultad:** medio (5). Implica cálculo automático y segmentación por canal.
 - **Complejidad:** medio (5). Requiere integrar costos y precios de múltiples fuentes.
 - **Riesgo:** bajo (3). Riesgo medio por variabilidad en la estructura de costos.
 - **Story Points** = $5 + 5 + 3 = 13$.
 - **Prioridad:** alta.
- **Historia de usuario 5:** como analista financiero, quiero integrar dinámicamente los costos de insumos y logística al cálculo de márgenes para reflejar los cambios del entorno.
 - **Dificultad:** medio (5). Requiere modelar funciones de costo con datos variables.
 - **Complejidad:** bajo (3). Debe incluir actualizaciones periódicas y automatización.
 - **Riesgo:** muy bajo (1). Los datos suelen estar disponibles en el ERP.
 - **Story Points** = $5 + 3 + 1 = 8$ (valor más próximo en la serie de Fibonacci).
 - **Prioridad:** media.

6.3. Épica 3: simulación de escenarios

- **Historia de usuario 6:** como director comercial, quiero simular escenarios ante cambios en el precio de venta de los productos, para anticipar su impacto económico.
 - **Dificultad:** medio (5). Requiere árboles de decisión y simulación dinámica.
 - **Complejidad:** medio (5). Debe considerar elasticidades y sensibilidad de variables.
 - **Riesgo:** bajo (3). El modelo puede ser sensible a inputs inciertos.
 - **Story Points** = $5 + 5 + 3 = 13$.
 - **Prioridad:** alta.
- **Historia de usuario 7:** como gerente de planta, quiero modelar la capacidad instalada y la ociosidad según distintos niveles de demanda, para optimizar la producción.
 - **Dificultad:** medio (5). Requiere modelar la planta con datos reales de capacidad.
 - **Complejidad:** muy bajo (1). Supone solo parametrizar restricciones operativas.
 - **Riesgo:** bajo (1). Baja incertidumbre si se dispone de datos internos.
 - **Story Points** = $5 + 1 + 1 = 8$ (valor próximo en la serie de Fibonacci).
 - **Prioridad:** media.
- **Historia de usuario 8 (opcional):** como analista, quiero simular restricciones operativas (por ejemplo, cortes de gas), para cuantificar su impacto en márgenes.
 - **Dificultad:** bajo (3). Se basa en inputs simples de disponibilidad.
 - **Complejidad:** muy bajo (1). La lógica de escenarios es sencilla.
 - **Riesgo:** muy bajo (1). No tiene efectos permanentes en el modelo.
 - **Story Points** = $3 + 1 + 1 = 5$.
 - **Prioridad:** baja.

6.4. Épica 4: visualización y prueba de concepto (PoC)

- **Historia de usuario 9:** como usuario final, quiero visualizar los resultados de la predicción y simulación en dashboards interactivos, para tomar decisiones más informadas.
 - **Dificultad:** medio (5). Requiere conectar datos y construir visualizaciones.
 - **Complejidad:** bajo (3). Necesita UX clara y efectiva.
 - **Riesgo:** bajo (1). Baja probabilidad de error si se usan herramientas estándar.
 - **Story Points** = $5 + 3 + 1 = 8$.
 - **Prioridad:** alta.
- **Historia de usuario 10:** como responsable de planificación, quiero comparar escenarios base y alternativos en un mismo tablero, para evaluar planes de acción.
 - **Dificultad:** bajo (3). Visualización comparativa.
 - **Complejidad:** bajo (3). Requiere lógica de selección de escenarios.
 - **Riesgo:** muy bajo (1). Los datos ya están procesados.
 - **Story Points** = $3 + 3 + 1 = 8$ (valor próximo en la serie de Fibonacci).
 - **Prioridad:** media.
- **Historia de usuario 11 (opcional):** como responsable financiero, quiero exportar la información a Excel desde Power BI o Streamlit, para facilitar reportes internos.
 - **Dificultad:** bajo (2). Se trata de una funcionalidad estándar.
 - **Complejidad:** muy bajo (1). Muy baja.
 - **Riesgo:** nulo (0). Prácticamente nulo.
 - **Story Points** = $2 + 1 + 0 = 3$.
 - **Prioridad:** baja.

7. Criterios de aceptación de historias de usuario

En esta sección se definen los criterios de aceptación para cada historia de usuario del *backlog*, con el objetivo de establecer condiciones específicas, medibles y verificables que permitan validar la funcionalidad desarrollada.

Se utiliza el modelo Gherkin, un enfoque ampliamente adoptado en metodologías ágiles para describir escenarios de prueba funcional en lenguaje natural estructurado. Este modelo sigue una sintaxis sencilla basada en las siguientes expresiones clave:

- **Dado (Given):** establece el contexto inicial o condiciones previas.
- **Cuando (When):** describe la acción que se va a realizar.
- **Entonces (Then):** indica el resultado esperado que permite validar el cumplimiento de la funcionalidad.

Este formato permite una comprensión clara tanto por parte de usuarios técnicos como no técnicos, facilitando la colaboración y alineamiento entre desarrolladores, analistas y *stakeholders* del proyecto.

7.1. Épica 1: modelado de demanda

Criterios de aceptación HU1: modelo de predicción de demanda semanal por producto

- **Criterio 1**
Dado un conjunto de datos históricos de ventas,
cuando se ejecuta el modelo de predicción,
entonces se obtiene una proyección semanal por producto con un error MAPE menor al 30 %.
- **Criterio 2**
Dado un conjunto de productos activos,
cuando el modelo es entrenado,
entonces debe generar una salida estructurada (tabla o archivo) con las predicciones para las siguientes 8 semanas.
- **Criterio 3**
Dado un modelo ya validado,
cuando se realiza una actualización mensual de datos,
entonces el modelo debe actualizar sus parámetros automáticamente y emitir un nuevo pronóstico.

Criterios de aceptación HU2: alertas ante desvíos en la demanda real vs. proyectada

- **Criterio 1**
Dado un desvío superior al 30 % entre la demanda real y la proyectada,
cuando se registre la venta semanal,
entonces el sistema debe generar una alerta visual o por correo electrónico.
- **Criterio 2**
Dado un conjunto de productos monitoreados,
cuando se detecten patrones de desvío sistemático en tres semanas consecutivas,
entonces se debe recomendar una revisión del modelo.

Criterios de aceptación HU3 (opcional): predicción agregada por línea de producto

- **Criterio 1**
Dado un grupo de productos pertenecientes a la misma línea,
cuando se ejecuta la predicción,
entonces se debe mostrar el resultado agregado en una única fila para la línea correspondiente.
- **Criterio 2**
Dado un usuario del área de producción,
cuando consulte la visualización,
entonces podrá seleccionar líneas de producto y visualizar la demanda proyectada semanalmente.

7.2. Épica 2: estimación de márgenes brutos por canal

Criterios de aceptación HU4: cálculo proyectado de márgenes brutos por producto y canal

- **Criterio 1**

Dado un conjunto de precios de venta estimados y costos variables actualizados,
cuando se ejecute el modelo de márgenes,
entonces debe calcularse automáticamente el margen bruto por producto y canal, expresado en valores absolutos y porcentuales.

- **Criterio 2**

Dado un conjunto de productos con diferentes estructuras de costos,
cuando se visualicen los resultados,
entonces el sistema debe permitir ordenar por margen y detectar productos con contribución negativa.

- **Criterio 3**

Dado un escenario base de estimaciones,
cuando se cambien los precios o costos,
entonces los nuevos márgenes deben actualizarse dinámicamente para reflejar el impacto de la variación.

Criterios de aceptación HU5: visualización de márgenes históricos y proyectados

- **Criterio 1**

Dado un producto seleccionado,
cuando se consulte su historial,
entonces deben mostrarse en una misma visualización sus márgenes brutos históricos y las proyecciones futuras.

- **Criterio 2**

Dado un conjunto de productos agrupados por canal,
cuando se visualicen sus márgenes,
entonces debe identificarse claramente cuáles tienen márgenes por debajo del umbral mínimo definido por la empresa.

7.3. Épica 3: simulación de escenarios operativos y financieros

Criterios de aceptación HU6: construcción de árbol de decisiones para escenarios

- **Criterio 1**

Dado un conjunto de decisiones estratégicas posibles (e.g., cambio de precio, capacidad instalada, compras),
cuando se cargan como insumos al sistema,
entonces el árbol de decisiones debe construir ramificaciones que representen los distintos caminos posibles y sus resultados económicos esperados.

- **Criterio 2**

Dado un árbol de decisiones generado,
cuando se selecciona una rama o nodo final,
entonces debe mostrarse la utilidad proyectada asociada y el escenario correspondiente de demanda y costos.

- **Criterio 3**

Dado un escenario base,
cuando se simula un cambio en una variable crítica (e.g., precio de insumo),
entonces el árbol debe recalcular los resultados y mostrar las variaciones comparadas al escenario original.

Criterios de aceptación HU7: análisis de sensibilidad sobre márgenes

- **Criterio 1**

Dado un conjunto de variables sensibles (precio, volumen, costos),
cuando se modifique su valor en un rango definido,
entonces el sistema debe mostrar en tiempo real cómo varía el margen bruto por producto y canal.

- **Criterio 2**

Dado un análisis de sensibilidad realizado,
cuando se identifique una variable que impacta fuertemente en el margen,
entonces el sistema debe destacarla como crítica mediante una etiqueta visual o comentario.

Criterios de aceptación HU8 (opcional): simulación de restricciones operativas

- **Criterio 1**

Dado un conjunto de restricciones operativas posibles (por ejemplo, cortes de gas o capacidad limitada),
cuando se seleccione una o varias restricciones en el sistema,
entonces debe simularse automáticamente el impacto de estas restricciones sobre el margen bruto total y por canal.

- **Criterio 2**

Dado un escenario con restricciones operativas aplicadas,
cuando se compare con el escenario base sin restricciones,
entonces el sistema debe mostrar claramente las diferencias de rendimiento económico y operatividad.

7.4. Épica 4: visualización y Prueba de Concepto (PoC)

Criterios de aceptación HU9: visualización de resultados en dashboards interactivos

- **Criterio 1**

Dado que el modelo de predicción y simulación ya generó resultados,
cuando se publiquen en una herramienta de visualización,
entonces el usuario debe poder consultar métricas clave por producto y canal de forma dinámica.

- **Criterio 2**

Dado un dashboard implementado en Power BI o Streamlit,
cuando el usuario interactúe con filtros de fecha o canal,
entonces los resultados deben actualizarse automáticamente reflejando la selección.

Criterios de aceptación HU10: comparación de escenarios

- **Criterio 1**

Dado un tablero que incluye datos de diferentes escenarios,
cuando el usuario seleccione un escenario base y uno alternativo,
entonces el sistema debe mostrar ambos en paralelo con sus diferencias destacadas.

- **Criterio 2**

Dado que se comparan márgenes y niveles de demanda entre escenarios,
cuando se utilicen herramientas visuales como gráficos de barras o tablas comparativas,
entonces el usuario debe poder identificar fácilmente el impacto de las decisiones evaluadas.

Criterios de aceptación HU11 (opcional): exportación a Excel

- **Criterio 1**

Dado que el usuario final accede al dashboard,
cuando seleccione una tabla de resultados,
entonces debe contar con una opción de exportar a Excel de forma directa y funcional.

- **Criterio 2**

Dado un archivo Excel exportado desde el sistema,
cuando el usuario lo abra,
entonces debe contener los datos visibles en pantalla con las mismas columnas y formatos.

8. Fases de CRISP-DM

8.1. Comprensión del negocio

El objetivo del proyecto es mejorar la planificación financiera en la familia de productos secos (salamines) del ciclo III de la empresa Cagnoli S.A., a través de un sistema inteligente que permita prever la demanda semanal, estimar márgenes brutos por canal y simular escenarios económicos-operativos. El valor agregado del uso de inteligencia artificial (IA) radica en reemplazar métodos heurísticos y subjetivos por herramientas objetivas, basadas en datos históricos y técnicas de *machine learning*. El éxito del proyecto será evaluado mediante mejoras en la precisión del pronóstico, la capacidad de análisis de escenarios, y el uso efectivo del sistema por parte de los equipos de planificación.

8.2. Comprensión de los datos

Los datos provienen de diversas fuentes internas, principalmente del sistema ERP de la compañía (Microsoft Dynamics), planillas de costos, inventarios, pedidos y reportes de ventas. Se cuenta con registros históricos de demanda semanal por producto, precios de venta, estructura de costos variables y datos de producción. La calidad de los datos es adecuada aunque presenta desafíos como series incompletas, variaciones de codificación, series con diferente dimensionalidad o a distinto nivel de detalle (granularidad).

8.3. Preparación de los datos

Durante esta fase se realizará la limpieza y transformación de los datos, lo cual incluye:

- Agregación de la demanda a nivel semanal por producto y canal.
- Tratamiento de valores faltantes y atípicos.
- Creación de nuevas variables explicativas (festividades, clima, tendencias, promociones).
- Normalización de series y construcción de *datasets* de entrenamiento y prueba.

La preparación de los datos es crítica para garantizar la calidad del modelo predictivo y la estabilidad de las simulaciones.

8.4. Modelado

El problema principal es de predicción de series temporales multivariadas y clasificación de escenarios. Se utilizarán modelos estadísticos y de aprendizaje automático como:

- SARIMA y Prophet, para capturar estacionalidad y tendencias.
- XGBoost y *Random Forest*, para incorporar múltiples variables predictoras y mejorar la robustez frente al ruido.
- Árboles de decisión y análisis de sensibilidad, para la simulación de escenarios y evaluación de impacto en márgenes.

8.5. Evaluación del modelo

Los modelos se evaluarán mediante métricas clásicas de rendimiento predictivo, tales como:

- MAE (*Mean Absolute Error*) y RMSE (*Root Mean Squared Error*) para pronóstico de demanda.
- Error absoluto medio en márgenes estimados para los resultados financieros.
- Validación cruzada con períodos de *backtesting* y análisis visual de ajuste de las predicciones.

8.6. Despliegue del modelo (opcional)

El modelo será implementado inicialmente como una prueba de concepto (PoC), mediante herramientas de visualización como Power BI o Streamlit. En esta instancia, el sistema funcionará como un tablero interno para los equipos de planificación, con la capacidad de simular diferentes escenarios y visualizar resultados de predicción. No se prevé integración con el ERP durante esta etapa inicial, aunque sí se contemplan recomendaciones para una futura integración en producción.

9. Desglose del trabajo en tareas

A continuación, se presenta el desglose técnico de las tareas correspondientes a las historias de usuario HU1 a HU4. Se estima una duración entre 2 y 8 horas por tarea, con una prioridad relativa asignada según su impacto en los entregables. Este desglose servirá como base para la planificación en las secciones siguientes (Diagrama de Gantt y Sprints).

Adicionalmente, al inicio se consideran las tareas de planificación del proyecto y selección del director, y finalmente se adicionan las tareas de preparación final de la documentación y práctica previo a la defensa del trabajo final.

9.0. Planificación del proyecto (Previo al desarrollo)

Historia de usuario	Tarea técnica	Estimación	Prioridad
N/A	Reunión inicial con el cliente y recolección de requerimientos	3 h	Alta
N/A	Lectura de documentos de contexto del negocio	2 h	Alta
N/A	Validación de problemas con stakeholders	3 h	Alta
N/A	Revisión de lineamientos académicos para el plan de proyecto	3 h	Alta
N/A	Búsqueda y contacto de potenciales directores	2 h	Media
N/A	Coordinación de reunión para validación del tema	3 h	Alta
N/A	Redacción inicial del documento de planificación	6 h	Alta
N/A	Elaboración del cronograma preliminar y entregables	3 h	Alta
N/A	Ajuste del documento según feedback	3 h	Alta
N/A	Preparación y entrega formal del plan	4 h	Alta

Cuadro 1. Desglose de tareas para la planificación del proyecto.

Total estimado parcial: 32 horas.

9.1. Épica 1: modelado de demanda

Historia de usuario	Tarea técnica	Estimación	Prioridad
HU1	Planificación del modelo de demanda	5 h	Alta
HU1	Selección de herramientas y estrategias de modelado	5 h	Alta
HU1	Análisis exploratorio de datos históricos	6 h	Alta
HU1	Visualización de estacionalidad y tendencias	6 h	Alta
HU1	Detección y tratamiento de outliers y valores nulos	6 h	Alta
HU1	Estandarización y generación de variables temporales	6 h	Alta
HU1	Configuración y entrenamiento inicial del modelo SARIMA	6 h	Alta
HU1	Ajuste fino de parámetros y tuning	6 h	Alta
HU1	Preparación de inputs y ejecución de Prophet	6 h	Alta
HU1	Ajuste y evaluación de estacionalidad y feriados	6 h	Alta
HU1	Optimización de hiperparámetros SARIMA y Prophet	6 h	Alta
HU1	Evaluación comparativa de precisión con métricas MAE y RMSE	6 h	Alta
HU1	Validación cruzada temporal con SARIMA y Prophet	6 h	Alta
HU1	Análisis de errores y comparación final entre modelos	6 h	Alta
HU1	Documentación técnica del modelo	8 h	Media
HU1	Pruebas con dataset de validación externo	5 h	Alta
HU1	Revisión de outputs y trazabilidad de resultados	5 h	Alta
HU1	Presentación de resultados a usuarios	8 h	Media

Cuadro 2. Desglose de tareas para HU1 (Épica 1: modelado de demanda.)

Historia de usuario	Tarea técnica	Estimación	Prioridad
HU2	Definición de variables objetivo y predictores / RF	5 h	Alta
HU2	Diseño de arquitectura y estrategia de modelado supervisado	5 h	Alta
HU2	Relevamiento de fuentes externas de datos (clima, estacionalidad)	5 h	Alta
HU2	Curado y validación de variables externas seleccionadas	5 h	Alta
HU2	Generación de nuevas variables explicativas	6 h	Alta
HU2	Evaluación de importancia relativa y reducción de dimensionalidad	6 h	Alta
HU2	Entrenamiento y validación de XG-Boost	6 h	Alta
HU2	Entrenamiento y validación de Random Forest	6 h	Alta
HU2	Comparación con SARIMA y Prophet en conjunto de test	5 h	Alta
HU2	Visualización de resultados y análisis de precisión	5 h	Alta
HU2	Visualización de errores y residual analysis	8 h	Media
HU2	Documentación de modelos y variables explicativas	8 h	Media
HU2	Reunión de feedback con equipo comercial	8 h	Media
HU3	Recolección de datos desde ERP de costos variables	5 h	Alta
HU3	Diseño de estructura de integración y procesamiento de costos	5 h	Alta
HU3	Modelado del costo variable por producto	5 h	Alta
HU3	Ajuste para variabilidad mensual, ajustes estacionales y otros costos	5 h	Alta
HU3	Integración con datos de precios de venta	8 h	Alta
HU3	Cálculo de márgenes proyectados	5 h	Alta
HU3	Comparativa de rentabilidad productos por canal	5 h	Alta
HU3	Análisis de sensibilidad de márgenes	8 h	Media
HU3	Validación con equipo financiero	8 h	Alta
HU3	Documentación de supuestos de cálculo	6 h	Media
HU3	Ajuste final del modelo y revisión	8 h	Media

Cuadro 3. Desglose de tareas para HU2 y HU3 (Épica 1: modelado de demanda).

Total estimado parcial HU1–HU3: 254 horas.

9.2. Épica 2: estimación de márgenes brutos por canal

Historia de usuario	Tarea técnica	Estimación	Prioridad
HU4	Análisis de escenarios críticos históricos	8 h	Alta
HU4	Diseño lógico de árbol de decisiones	8 h	Alta
HU4	Diseño del motor de simulación de escenarios	6 h	Alta
HU4	Codificación del motor de simulación de escenarios	6 h	Alta
HU4	Implementación de pruebas con cambios en precios	5 h	Alta
HU4	Implementación de pruebas con cambios en volúmenes	5 h	Alta
HU4	Visualización del impacto en márgenes y ventas	8 h	Media
HU4	Validación de escenarios con dirección	6 h	Media
HU4	Documentación de escenarios y recomendaciones	6 h	Media
HU4	Ajustes finales tras feedback del usuario	6 h	Media
HU5	Revisión bibliográfica de métodos de simulación	8 h	Alta
HU5	Diseño de arquitectura funcional del módulo de escenarios	6 h	Alta
HU5	Diseño de arquitectura técnica del módulo de escenarios	6 h	Alta
HU5	Lógica de escenarios base: codificación y pruebas	8 h	Alta
HU5	Lógica de escenarios alternativos: codificación y pruebas	8 h	Alta
HU5	Integración con outputs de modelos de predicción: análisis de compatibilidad	5 h	Alta
HU5	Integración con outputs de modelos de predicción: desarrollo e implementación	6 h	Alta
HU5	Validación preliminar de resultados del simulador	5 h	Alta
HU5	Validación de casos de uso con usuarios funcionales	5 h	Alta
HU5	Documentación funcional y técnica del simulador	8 h	Media
HU5	Revisión y ajuste en base a feedback	6 h	Media

Cuadro 4. Desglose de tareas para HU4 a HU5 (Épica 2: estimación de márgenes brutos por canal).

Total estimado HU4-HU5: 135 horas.

9.3. Épica 3: simulación de escenarios operativos y financieros

Historia de usuario	Tarea técnica	Estimación	Prioridad
HU6	Recolección de eventos exógenos históricos (clima, insumos)	5 h	Alta
HU6	Limpieza y estandarización de eventos exógenos	5 h	Alta
HU6	Modelado de impactos: análisis exploratorio	6 h	Alta
HU6	Modelado de impactos: implementación	6 h	Alta
HU6	Diseño de lógica para incorporar shocks al simulador	6 h	Alta
HU6	Codificación y pruebas de integración de shocks	6 h	Alta
HU6	Definición de casos de validación	5 h	Media
HU6	Contraste de resultados del modelo vs. escenarios reales	5 h	Media
HU6	Documentación de implementación y casos simulados	8 h	Media
HU7	Diseño de interfaz para inputs de simulación	5 h	Alta
HU7	Prototipado y revisión de interfaz con usuarios	5 h	Alta
HU7	Desarrollo de backend para procesamiento dinámico (parte 1)	6 h	Alta
HU7	Desarrollo de backend para procesamiento dinámico (parte 2)	6 h	Alta
HU7	Testeo de inputs y validación de comportamiento	8 h	Alta
HU7	Manual de uso interactivo del simulador	6 h	Media
HU7	Recopilación de feedback de usuarios funcionales	3 h	Media
HU7	Aplicación de mejoras iterativas al simulador	3 h	Media
HU8 (opcional)	Simulación de restricciones operativas críticas (fase 1)	5 h	Baja
HU8 (opcional)	Simulación de restricciones operativas críticas (fase 2)	5 h	Baja
HU8 (opcional)	Análisis del impacto financiero (datos y supuestos)	5 h	Baja
HU8 (opcional)	Cálculo de impacto en márgenes	5 h	Baja
HU8 (opcional)	Visualización de resultados y diseño de casos	8 h	Baja
HU8 (opcional)	Documentación breve del caso operativo simulado	4 h	Baja

Cuadro 5. Desglose de tareas para HU6 a HU8 (Épica 3: simulación de escenarios operativos y financieros).

Total estimado HU6-HU8: 126 horas.

9.4. Épica 4: visualización y Prueba de Concepto (PoC)

Historia de usuario	Tarea técnica	Estimación	Prioridad
HU9	Diseño inicial del layout para visualización de predicciones	5 h	Alta
HU9	Diseño del layout de márgenes y filtros interactivos	5 h	Alta
HU9	Desarrollo de dashboards: estructura y navegación	7 h	Alta
HU9	Desarrollo de dashboards: integración con librerías y UX	7 h	Alta
HU9	Conexión a fuentes de datos internas del modelo	5 h	Alta
HU9	Validación de conexión y pruebas de actualización automática	5 h	Alta
HU9	Validación de visualizaciones con usuarios clave	8 h	Alta
HU9	Ajuste de gráficos y filtros según feedback	6 h	Media
HU9	Redacción de manual de uso de dashboards	2 h	Media
HU9	Inclusión de ejemplos y pantallas en la documentación	2 h	Media
HU10	Diseño de panel comparativo de escenarios	8 h	Media
HU10	Lógica de selección de escenarios para comparación	8 h	Media
HU10	Implementación visual del comparador y controles	6 h	Media
HU10	Pruebas de casos base vs. alternativos	8 h	Media
HU10	Documentación técnica del panel comparador	3 h	Media
HU10	Documentación funcional con ejemplos de uso	3 h	Media
HU11 (opcional)	Desarrollo de exportación a Excel desde dashboard	4 h	Baja
HU11 (opcional)	Validación de formato y campos exportables	2 h	Baja
HU11 (opcional)	Documentación breve de exportación y soporte	2 h	Baja

Cuadro 6. Desglose de tareas para HU9 a HU11 (Épica 4: visualización y Prueba de Concepto (PoC)).

Total estimado parcial HU9–HU11: 96 horas.

9.5. Documentación y presentación final

Historia de usuario	Tarea técnica	Estimación	Prioridad
N/A	Revisión general del documento final y estructura	4 h	Alta
N/A	Redacción del resumen ejecutivo	3 h	Alta
N/A	Edición final y generación de versión PDF para entrega	3 h	Alta
N/A	Redacción de memoria (parte 1)	8 h	Alta
N/A	Redacción de memoria (parte 2)	8 h	Alta
N/A	Preparación de defensa	8 h	Alta

Cuadro 7. Desglose de tareas para la documentación y presentación final.

Total estimado parcial: 34 horas.

Estimación del costo total del proyecto:

Dado que el presente trabajo final requiere una dedicación estimada de 677 horas, y considerando un valor de referencia de USD 40 por hora de ingeniería, el costo total estimado del proyecto asciende a:

$$677 \text{ h} \times \text{USD } 40/\text{h} = \text{USD } 27.080$$

Esta estimación no representa un costo real incurrido, sino una valorización técnica del esfuerzo y dedicación requeridos para su ejecución, basada en estándares del mercado para profesionales especializados en inteligencia artificial y planificación financiera.

.....

Cantidad total de horas: (677 h)

.....

10. Diagrama de Gantt

El presente diagrama de Gantt representa de forma visual y cronológica todas las tareas planificadas para el desarrollo del trabajo final. La planificación contempla un total de aproximadamente 677 horas, distribuidas entre tareas técnicas (modelado, simulación, desarrollo de dashboards, validaciones, etc.) y tareas no técnicas (planificación, documentación, redacción de entregables y preparación de la presentación final).

Como primer paso en la planificación, se realizó una distribución detallada de tareas en 15 sprints quincenales. Estos sprints comenzarán luego de la fase inicial de planificación (29/04/2025 al 13/06/2025). Esta división en sprints permitió agrupar tareas según su prioridad, orden lógico y dependencias técnicas, con el fin de asegurar una carga equilibrada de entre 45 y 50 horas

por sprint. A partir de esta organización, se construyó el cronograma general representado en el diagrama de Gantt que se muestra a continuación.

El diagrama está estructurado con las tareas organizadas verticalmente y el tiempo representado en el eje horizontal. Cada tarea se asocia a un sprint específico, y se ha utilizado una codificación de colores para distinguir entre tareas técnicas (azul) y no técnicas (naranja). Además, se han configurado las dependencias entre tareas para reflejar adecuadamente la lógica del flujo de trabajo, con el fin de evitar solapamientos innecesarios y permitiendo la trazabilidad del avance del proyecto.

Este cronograma inicia con las actividades de planificación del proyecto (Etapa 0) y culmina con la entrega final y preparación de la defensa, hacia mediados de enero de 2026. En las siguientes figuras se presenta el diagrama de Gantt del proyecto por partes:



Figura 2. Diagrama de Gantt del Proyecto - Parte 1 de 4.

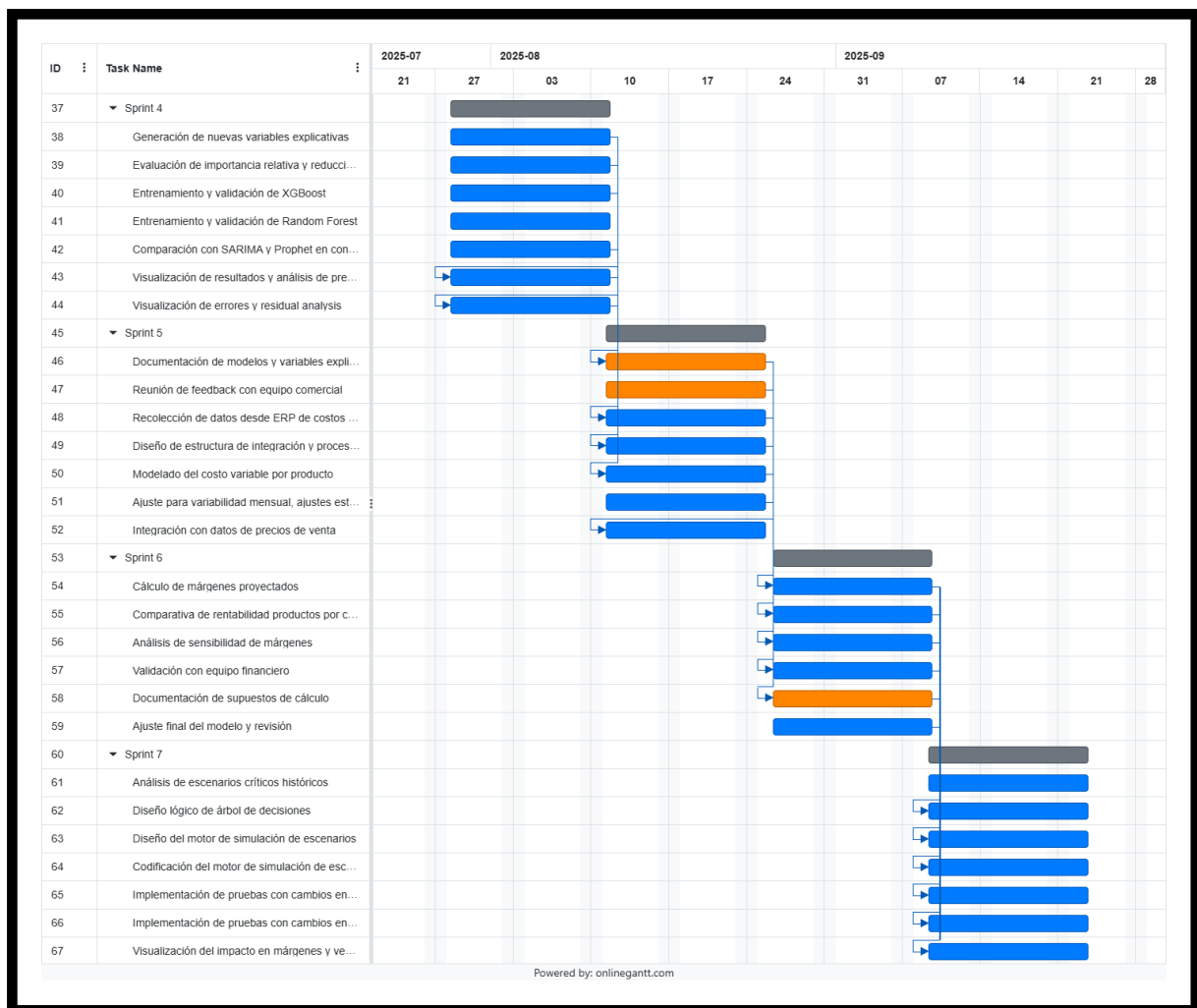


Figura 3. Diagrama de Gantt del Proyecto - Parte 2 de 4.



Figura 4. Diagrama de Gantt del Proyecto - Parte 3 de 4.



Figura 5. Diagrama de Gantt del Proyecto - Parte 4 de 4.

11. Planificación de Sprints

El objetivo de esta sección es organizar la ejecución del proyecto en sprints de trabajo con una distribución equilibrada de la carga horaria. La planificación abarca un total de aproximadamente 600 horas, de las cuales entre 480 y 500 horas corresponden a tareas técnicas, y entre 100 y 120 horas a tareas no técnicas, como planificación, redacción de la memoria y preparación de la defensa.

Los sprints definidos tienen una duración de dos semanas cada uno. El primer sprint comienza el 16 de junio de 2025, tras finalizar la etapa de planificación. En total se estructuraron 15 sprints, de los cuales se considera que el último estará destinado exclusivamente a las tareas de cierre del proyecto. La asignación de tareas por sprint se realizó a partir del desglose previo de historias de usuario y tareas técnicas, respetando un límite de entre 45 y 50 horas por sprint.

La tabla incluida en esta sección muestra cada sprint junto con las tareas que lo componen, el número de horas estimadas, el responsable (el autor del proyecto) y el porcentaje de avance previsto o completado. Esta planificación permite establecer prioridades, realizar un seguimiento riguroso del progreso y facilitar instancias de revisión periódica.

Cuadro 8. Planificación inicial (previo al comienzo de sprints).

Sprint	HU o fase	Tarea	Horas	Responsable	% Completado
N/A	N/A	Reunión inicial con el cliente y recolección de requerimientos	3 h	Alumno	100 %
N/A	N/A	Lectura de documentos de contexto del negocio	2 h	Alumno	100 %
N/A	N/A	Validación de problemas con stakeholders	3 h	Alumno	100 %
N/A	N/A	Revisión de lineamientos académicos para el plan de proyecto	3 h	Alumno	100 %
N/A	N/A	Búsqueda y contacto de potenciales directores	2 h	Alumno	20 %
N/A	N/A	Coordinación de reunión para validación del tema	3 h	Alumno	0 %
N/A	N/A	Redacción inicial del documento de planificación	6 h	Alumno	50 %
N/A	N/A	Elaboración del cronograma preliminar y entregables	3 h	Alumno	50 %
N/A	N/A	Ajuste del documento según feedback	3 h	Alumno	30 %
N/A	N/A	Preparación y entrega formal del plan	4 h	Alumno	0 %

Cuadro 9. Planificación de Sprints 1 a 2.

Sprint	HU o fase	Tarea	Horas	Responsable	% Completado
Sprint 1	HU1	Planificación del modelo de demanda	5 h	Alumno	0 %
Sprint 1	HU1	Selección de herramientas y estrategias de modelado	5 h	Alumno	0 %
Sprint 1	HU1	Análisis exploratorio de datos históricos	6 h	Alumno	0 %
Sprint 1	HU1	Visualización de estacionalidad y tendencias	6 h	Alumno	0 %
Sprint 1	HU1	Detección y tratamiento de outliers y valores nulos	6 h	Alumno	0 %
Sprint 1	HU1	Estandarización y generación de variables temporales	6 h	Alumno	0 %
Sprint 1	HU1	Configuración y entrenamiento inicial del modelo SARIMA	6 h	Alumno	0 %
Sprint 2	HU1	Ajuste fino de parámetros y tuning	6 h	Alumno	0 %
Sprint 2	HU1	Preparación de inputs y ejecución de Prophet	6 h	Alumno	0 %
Sprint 2	HU1	Ajuste y evaluación de estacionalidad y feriados	6 h	Alumno	0 %
Sprint 2	HU1	Optimización de hiperparámetros SARIMA y Prophet	6 h	Alumno	0 %
Sprint 2	HU1	Evaluación comparativa de precisión con métricas MAE y RMSE	6 h	Alumno	0 %
Sprint 2	HU1	Validación cruzada temporal con SARIMA y Prophet	6 h	Alumno	0 %
Sprint 2	HU1	Análisis de errores y comparación final entre modelos	6 h	Alumno	0 %

Cuadro 10. Planificación de Sprints 3 a 4.

Sprint	HU o fase	Tarea	Horas	Responsable	% Completado
Sprint 3	HU1	Documentación técnica del modelo	8 h	Alumno	0 %
Sprint 3	HU1	Pruebas con dataset de validación externo	5 h	Alumno	0 %
Sprint 3	HU1	Revisión de outputs y trazabilidad de resultados	5 h	Alumno	0 %
Sprint 3	HU1	Presentación de resultados a usuarios	8 h	Alumno	0 %
Sprint 3	HU1	Definición de variables objetivo y predictores / RF	5 h	Alumno	0 %
Sprint 3	HU1	Diseño de arquitectura y estrategia de modelado supervisado	5 h	Alumno	0 %
Sprint 3	HU1	Relevamiento de fuentes externas de datos (clima, estacionalidad)	5 h	Alumno	0 %
Sprint 3	HU1	Curado y validación de variables externas seleccionadas	5 h	Alumno	0 %
Sprint 4	HU2	Generación de nuevas variables explicativas	6 h	Alumno	0 %
Sprint 4	HU2	Evaluación de importancia relativa y reducción de dimensionalidad	6 h	Alumno	0 %
Sprint 4	HU2	Entrenamiento y validación de XGBoost	6 h	Alumno	0 %
Sprint 4	HU2	Entrenamiento y validación de Random Forest	6 h	Alumno	0 %
Sprint 4	HU2	Comparación con SARIMA y Prophet en conjunto de test	5 h	Alumno	0 %
Sprint 4	HU2	Visualización de resultados y análisis de precisión	5 h	Alumno	0 %
Sprint 4	HU2	Visualización de errores y <i>residual analysis</i>	8 h	Alumno	0 %

Cuadro 11. Planificación de Sprints 5 a 6.

Sprint	HU o fase	Tarea	Horas	Responsable	% Completado
Sprint	HU o fase	Tarea	Horas	Responsable	% Completado
Sprint 5	HU2	Documentación de modelos y variables explicativas	8 h	Alumno	0 %
Sprint 5	HU2	Reunión de feedback con equipo comercial	8 h	Alumno	0 %
Sprint 5	HU3	Recolección de datos desde ERP de costos variables	5 h	Alumno	0 %
Sprint 5	HU3	Diseño de estructura de integración y procesamiento de costos	5 h	Alumno	0 %
Sprint 5	HU3	Modelado del costo variable por producto	5 h	Alumno	0 %
Sprint 5	HU3	Ajuste para variabilidad mensual, ajustes estacionales y otros costos	5 h	Alumno	0 %
Sprint 5	HU3	Integración con datos de precios de venta	8 h	Alumno	0 %
Sprint 6	HU3	Cálculo de márgenes proyectados	5 h	Alumno	0 %
Sprint 6	HU3	Comparativa de rentabilidad productos por canal	5 h	Alumno	0 %
Sprint 6	HU3	Análisis de sensibilidad de márgenes	8 h	Alumno	0 %
Sprint 6	HU3	Validación con equipo financiero	8 h	Alumno	0 %
Sprint 6	HU3	Documentación de supuestos de cálculo	6 h	Alumno	0 %
Sprint 6	HU3	Ajuste final del modelo y revisión	8 h	Alumno	0 %

Cuadro 12. Planificación de Sprints 7 a 8.

Sprint	HU o fase	Tarea	Horas	Responsable	% Completado
Sprint 7	HU4	Análisis de escenarios críticos históricos	8 h	Alumno	0 %
Sprint 7	HU4	Diseño lógico de árbol de decisiones	8 h	Alumno	0 %
Sprint 7	HU4	Diseño del motor de simulación de escenarios	6 h	Alumno	0 %
Sprint 7	HU4	Codificación del motor de simulación de escenarios	6 h	Alumno	0 %
Sprint 7	HU4	Implementación de pruebas con cambios en precios	5 h	Alumno	0 %
Sprint 7	HU4	Implementación de pruebas con cambios en volúmenes	5 h	Alumno	0 %
Sprint 7	HU4	Visualización del impacto en márgenes y ventas	8 h	Alumno	0 %
Sprint	HU o fase	Tarea	Horas	Responsable	% Completado
Sprint 8	HU4	Validación de escenarios con dirección	6 h	Alumno	0 %
Sprint 8	HU4	Documentación de escenarios y recomendaciones	6 h	Alumno	0 %
Sprint 8	HU4	Ajustes finales tras feedback del usuario	6 h	Alumno	0 %
Sprint 8	HU5	Revisión bibliográfica de métodos de simulación	8 h	Alumno	0 %
Sprint 8	HU5	Diseño de arquitectura funcional del módulo de escenarios	6 h	Alumno	0 %
Sprint 8	HU5	Diseño de arquitectura técnica del módulo de escenarios	6 h	Alumno	0 %
Sprint 8	HU5	Lógica de escenarios base: codificación y pruebas	8 h	Alumno	0 %

Cuadro 13. Planificación de Sprints 9 a 10.

Sprint	HU o fase	Tarea	Horas	Responsable	% Completado
Sprint 9	HU5	Lógica de escenarios alternativos: codificación y pruebas	8 h	Alumno	0 %
Sprint 9	HU5	Integración con outputs de modelos de predicción: análisis de compatibilidad	5 h	Alumno	0 %
Sprint 9	HU5	Integración con outputs de modelos de predicción: desarrollo e implementación	6 h	Alumno	0 %
Sprint 9	HU5	Validación preliminar de resultados del simulador	5 h	Alumno	0 %
Sprint 9	HU5	Validación de casos de uso con usuarios funcionales	5 h	Alumno	0 %
Sprint 9	HU5	Documentación funcional y técnica del simulador	8 h	Alumno	0 %
Sprint 9	HU5	Revisión y ajuste en base a feedback	6 h	Alumno	0 %
Sprint	HU o fase	Tarea	Horas	Responsable	% Completado
Sprint 10	HU6	Recolección de eventos exógenos históricos (clima, insumos)	5 h	Alumno	0 %
Sprint 10	HU6	Limpieza y estandarización de eventos exógenos	5 h	Alumno	0 %
Sprint 10	HU6	Modelado de impactos: análisis exploratorio	6 h	Alumno	0 %
Sprint 10	HU6	Modelado de impactos: implementación	6 h	Alumno	0 %
Sprint 10	HU6	Diseño de lógica para incorporar shocks al simulador	6 h	Alumno	0 %
Sprint 10	HU6	Codificación y pruebas de integración de shocks	6 h	Alumno	0 %
Sprint 10	HU6	Definición de casos de validación	5 h	Alumno	0 %
Sprint 10	HU6	Contraste de resultados del modelo vs. escenarios reales	5 h	Alumno	0 %

Cuadro 14. Planificación de Sprints 11 a 12.

Sprint	HU o fase	Tarea	Horas	Responsable	% Completado
Sprint 11	HU6	Documentación de implementación y casos simulados	8 h	Alumno	0 %
Sprint 11	HU7	Diseño de interfaz para inputs de simulación	5 h	Alumno	0 %
Sprint 11	HU7	Prototipado y revisión de interfaz con usuarios	5 h	Alumno	0 %
Sprint 11	HU7	Desarrollo de backend para procesamiento dinámico (parte 1)	6 h	Alumno	0 %
Sprint 11	HU7	Desarrollo de backend para procesamiento dinámico (parte 2)	6 h	Alumno	0 %
Sprint 11	HU7	Testeo de inputs y validación de comportamiento	8 h	Alumno	0 %
Sprint 11	HU7	Manual de uso interactivo del simulador	6 h	Alumno	0 %
Sprint 11	HU7	Recopilación de feedback de usuarios funcionales	3 h	Alumno	0 %
Sprint 11	HU7	Aplicación de mejoras iterativas al simulador	3 h	Alumno	0 %
Sprint 12	HU8	Simulación de restricciones operativas críticas (fase 1)	5 h	Alumno	0 %
Sprint 12	HU8	Simulación de restricciones operativas críticas (fase 2)	5 h	Alumno	0 %
Sprint 12	HU8	Análisis del impacto financiero (datos y supuestos)	5 h	Alumno	0 %
Sprint 12	HU8	Cálculo de impacto en márgenes	5 h	Alumno	0 %
Sprint 12	HU8	Visualización de resultados y diseño de casos	8 h	Alumno	0 %
Sprint 12	HU8	Documentación breve del caso operativo simulado	4 h	Alumno	0 %
Sprint 12	HU9	Diseño inicial del layout para visualización de predicciones	5 h	Alumno	0 %
Sprint 12	HU9	Diseño del layout de márgenes y filtros interactivos	5 h	Alumno	0 %

Cuadro 15. Planificación de Sprints 13 a 14.

Sprint	HU o fase	Tarea	Horas	Responsable	% Completado
Sprint 13	HU9	Desarrollo de dashboards: estructura y navegación	7 h	Alumno	0 %
Sprint 13	HU9	Desarrollo de dashboards: integración con librerías y UX	7 h	Alumno	0 %
Sprint 13	HU9	Conexión a fuentes de datos internas del modelo	5 h	Alumno	0 %
Sprint 13	HU9	Validación de conexión y pruebas de actualización automática	5 h	Alumno	0 %
Sprint 13	HU9	Validación de visualizaciones con usuarios clave	8 h	Alumno	0 %
Sprint 13	HU9	Ajuste de gráficos y filtros según feedback	6 h	Alumno	0 %
Sprint 13	HU9	Redacción de manual de uso de dashboards	2 h	Alumno	0 %
Sprint 13	HU9	Inclusión de ejemplos y pantallas en la documentación	2 h	Alumno	0 %
Sprint 14	HU10	Diseño de panel comparativo de escenarios	8 h	Alumno	0 %
Sprint 14	HU10	Lógica de selección de escenarios para comparación	8 h	Alumno	0 %
Sprint 14	HU10	Implementación visual del comparador y controles	6 h	Alumno	0 %
Sprint 14	HU10	Pruebas de casos base vs. alternativos	8 h	Alumno	0 %
Sprint 14	HU10	Documentación técnica del panel comparador	3 h	Alumno	0 %
Sprint 14	HU10	Documentación funcional con ejemplos de uso	3 h	Alumno	0 %
Sprint 14	HU11	Desarrollo de exportación a Excel desde dashboard	4 h	Alumno	0 %
Sprint 14	HU11	Validación de formato y campos exportables	2 h	Alumno	0 %
Sprint 14	HU11	Documentación breve de exportación y soporte	2 h	Alumno	0 %

Cuadro 16. Planificación de Sprint 15.

Sprint	HU o fase	Tarea	Horas	Responsable	% Completado
Sprint 15	N/A	Revisión general del documento final y estructura	4 h	Alumno	0 %
Sprint 15	N/A	Redacción del resumen ejecutivo	3 h	Alumno	0 %
Sprint 15	N/A	Edición final y generación de versión PDF para entrega	3 h	Alumno	0 %
Sprint 15	N/A	Redacción de memoria (parte 1)	8 h	Alumno	0 %
Sprint 15	N/A	Redacción de memoria (parte 2)	8 h	Alumno	0 %
Sprint 15	N/A	Preparación de defensa	8 h	Alumno	0 %

12. Normativa y cumplimiento de datos (gobernanza)

El presente proyecto se apoya en el uso de datos operativos, comerciales y financieros provenientes de una empresa privada del sector alimenticio, específicamente vinculada a la producción y comercialización de fiambres y chacinados. Estos datos incluyen volúmenes de producción, series históricas de demanda, precios, márgenes, eventos exógenos y estructuras de costos, entre otros elementos. Dado el carácter sensible y estratégico de parte de esta información, resulta clave garantizar el cumplimiento normativo y el respeto de los principios de gobernanza de datos a lo largo de todo el desarrollo del proyecto.

12.1. Normativas aplicables

Los datos utilizados no incluyen información personal de clientes, empleados ni terceros identificables, por lo tanto no se encuentran comprendidos bajo la Ley Nacional 25.326 de Protección de Datos Personales (Argentina), ni bajo el Reglamento General de Protección de Datos de la Unión Europea (GDPR), ni normativas específicas como HIPAA (para el ámbito de la salud). Sin embargo, al trabajar con datos de carácter confidencial, operativos y estratégicos de una empresa, deben observarse los siguientes principios normativos y éticos:

- **Confidencialidad:** todo el procesamiento y análisis de datos se realiza bajo un acuerdo explícito con la empresa, garantizando la confidencialidad de la información.
- **Licitud y finalidad:** los datos son utilizados exclusivamente para los fines establecidos en el marco de este proyecto académico y no serán reutilizados con otros propósitos sin el consentimiento explícito de la empresa.
- **Minimización de datos:** se evita recolectar información innecesaria, y se emplea únicamente aquella requerida para alcanzar los objetivos del proyecto.
- **Transparencia:** se documenta el origen, tratamiento y aplicación de cada conjunto de datos, lo que facilita auditorías internas o académicas.

12.2. Condiciones de uso de los datos

El acceso a los datos fue autorizado por la organización en el marco de un proyecto colaborativo, donde el autor del presente trabajo se desempeña como consultor externo. Esta relación garantiza un marco de trabajo ético y profesional, bajo los siguientes términos:

- Acceso controlado: los datos no fueron compartidos públicamente ni almacenados en plataformas abiertas. El almacenamiento y procesamiento se realizará en entornos locales o corporativos seguros.
- No publicación sin anonimización: en caso de presentaciones académicas o publicaciones, cualquier dato será anonimizado o presentado de forma agregada, de modo de preservar la confidencialidad del negocio.
- Sin involucramiento de datos personales: no se procesan nombres, identificadores personales, correos electrónicos ni ningún tipo de dato sensible sobre personas físicas.
- Consentimiento institucional: el uso de los datos fue aprobado por los responsables de la empresa con el objetivo de contribuir al desarrollo del sistema de planificación y control.

12.3. Evaluación de restricciones y viabilidad

Desde el punto de vista legal y técnico, el uso de los datos es viable, siempre que se mantengan los criterios establecidos anteriormente. No existen restricciones contractuales, regulatorias o legales explícitas que impidan su tratamiento para fines académicos internos. Tampoco se requiere consentimiento de terceros, dado que no se trata de información asociada a individuos.

No obstante, se reconoce que parte de la información tiene valor estratégico para la empresa (e.g., márgenes, simulaciones de precios, estructura de costos), por lo cual se reafirma el compromiso de mantener la confidencialidad y de evitar cualquier exposición no autorizada de los resultados.

12.4. Gobernanza de datos

El proyecto promueve buenas prácticas de gobernanza de datos. Para lograr esto se asegura lo siguiente:

- Registro detallado de las fuentes de datos utilizadas, sus características, y las transformaciones aplicadas.
- Reproducibilidad del análisis: todos los procesos pueden ser auditados internamente por la empresa.
- Integración con políticas internas de seguridad: el diseño del modelo considera que, en etapas futuras, podría integrarse con herramientas empresariales como Power BI, Streamlit o el ERP Microsoft Dynamics, bajo protocolos seguros.
- Documentación y versionado: cada etapa de desarrollo ha sido documentada para facilitar el traspaso y sostenibilidad del sistema en el tiempo.

Conclusión

El presente proyecto cumple con los lineamientos legales, éticos y técnicos en materia de uso de datos. Si bien no se manejan datos personales ni se viola ninguna normativa de protección de datos vigente, se aplican estándares de gobernanza orientados a garantizar la transparencia, trazabilidad y confidencialidad de la información utilizada. Esto no solo resguarda a la organización involucrada, sino que también fortalece la robustez y profesionalismo del trabajo realizado.

13. Gestión de riesgos

En esta sección se identifican y analizan los principales riesgos asociados al desarrollo del presente proyecto. Se evalúan tanto su severidad (S) como su probabilidad de ocurrencia (O), utilizando una escala del 1 al 10, y se calcula el RPN (Risk Priority Number) como producto de ambas variables. A continuación, se presentan las justificaciones, la tabla de riesgos y los planes de mitigación correspondientes.

a) Identificación de los riesgos (al menos cinco) y estimación de sus consecuencias:

Riesgo 1: retrasos en la entrega de datos por parte del área financiera o comercial.

- Severidad (S): 8. Los datos son necesarios para alimentar los modelos de márgenes y simulación, afectando el cronograma completo si no están disponibles.
- Probabilidad de ocurrencia (O): 5. Aunque existe buena predisposición, pueden ocurrir demoras por prioridades internas o validaciones.

Riesgo 2: incompatibilidad entre formatos de datos extraídos del ERP y lo requerido por el modelo.

- Severidad (S): 6. Puede requerir tiempo adicional en preprocesamiento y scripts de transformación.
- Ocurrencia (O): 6. Ya se observaron casos de formatos inconsistentes entre áreas.

Riesgo 3: incertidumbre en el comportamiento de modelos de IA ante cambios estacionales abruptos.

- Severidad (S): 7. Podría reducir la confiabilidad del modelo de predicción de demanda.
- Ocurrencia (O): 6. Históricamente hubo saltos estacionales no triviales en la demanda.

Riesgo 4: falta de disponibilidad del consultor (responsable del proyecto) por compromisos laborales.

- Severidad (S): 8. El proyecto depende fuertemente del responsable.

- Ocurrencia (O): 4. Aunque el calendario está planificado, existe carga de trabajo simultánea.

Riesgo 5: dificultad para validar escenarios simulados con usuarios clave por falta de participación.

- Severidad (S): 6. Los escenarios pueden no ser creíbles ni útiles si no se validan con expertos.
- Ocurrencia (O): 5. Ya se experimentaron demoras en validaciones anteriores.

Riesgo 6: fallos técnicos en ejecución de modelos (capacidad limitada en entorno local).

- Severidad (S): 9. Puede frenar completamente tareas de entrenamiento o simulación.
- Ocurrencia (O): 4. El entorno local tiene recursos limitados, pero se prevé migración a la nube si es necesario.

Riesgo 7: dificultad para integrar márgenes y demanda en simulación en el tiempo proyectado.

- Severidad (S): 7. Puede dejar incompleta una de las funcionalidades principales del trabajo.
- Ocurrencia (O): 5. La integración requiere diseño modular y sincronización entre subsistemas.

Riesgo 8: rechazo del modelo final por usuarios por falta de explicabilidad.

- Severidad (S): 6. Disminuye su aplicabilidad y genera resistencia a su uso.
- Ocurrencia (O): 3. Dependerá de cómo se comuniquen los resultados y del nivel de documentación.

Riesgo 9: pérdida de avances por falta de backups.

- Severidad (S): 7. Implicaría rehacer parte del trabajo.
- Ocurrencia (O): 2. Se hacen respaldos, pero podrían no estar actualizados.

Riesgo 10: cambios en prioridades internas de la empresa que posterguen validaciones.

- Severidad (S): 6. Afecta entregables planificados.
- Ocurrencia (O): 4. Aunque el proyecto está acordado, pueden cambiar los tiempos según la coyuntura.

b) Tabla de gestión de riesgos: (El RPN se calcula como $RPN=S \times O$)

Riesgo	S	O	RPN	S*	O*	RPN*
1. Retraso en entrega de datos	8	5	40	6	3	18
2. Incompatibilidad de formatos de datos	6	6	36	4	4	16
3. Modelos no robustos a cambios estacionales	7	6	42	5	4	20
4. Falta de disponibilidad del responsable	8	4	32	-	-	-
5. Escasa participación en validación de escenarios	6	5	30	-	-	-
6. Fallos técnicos en entorno local	9	4	36	5	3	15
7. Problemas en integración márgenes-demanda	7	5	35	5	3	15
8. Rechazo por falta de explicabilidad	6	3	18	-	-	-
9. Pérdida de avances sin backup	7	2	14	-	-	-
10. Cambio de prioridades internas	6	4	24	-	-	-

Cuadro 17. Tabla de gestión de riesgos (se considera crítica toda RPN mayor o igual a 35).

Criterio adoptado:

Se tomarán medidas de mitigación en los riesgos cuyos números de RPN sean mayores o iguales a 35

Nota: los valores marcados con (*) en la tabla corresponden luego de haber aplicado la mitigación.

c) Plan de mitigación de los riesgos que originalmente excedían el RPN máximo establecido:

Dado que varios de los riesgos identificados comparten factores de origen comunes, (como la dependencia de datos externos, limitaciones técnicas y plazos acotados), se adopta una estrategia general de mitigación centrada en: (I) validación temprana, (II) automatización de tareas críticas, (III) coordinación proactiva con áreas responsables y (IV) planificación de escenarios alternativos. Esta estrategia se adapta específicamente a cada riesgo para reducir su severidad y/o probabilidad de ocurrencia.

Riesgo 1: retrasos en la entrega de datos por parte del área financiera o comercial.

Plan de mitigación específico: coordinar reuniones tempranas para alinear expectativas de entrega, generar alertas automáticas ante demoras y contar con fuentes de datos de respaldo.

- Severidad (S): 6. Se reduce mediante una mayor capacidad de reacción frente a demoras previstas.
- Ocurrencia (O): 3. Se reduce al establecer canales de comunicación y fechas acordadas formalmente.

Riesgo 2: incompatibilidad entre formatos de datos extraídos del ERP y lo requerido por el modelo

Plan de mitigación específico: desarrollar scripts de conversión automatizados y definir plantillas estándar para el intercambio de datos desde el inicio del proyecto.

- Severidad (S^*): 4. Se reduce al evitar errores manuales y acelerar la adaptación de datos.
- Probabilidad de ocurrencia (O^*): 4. Disminuye con la validación temprana de estructuras de datos.

Riesgo 3: incertidumbre en el comportamiento de modelos de IA ante cambios estacionales abruptos

Plan de mitigación específico: aplicar validación cruzada temporal, incorporar variables exógenas (clima, calendario), y mantener actualizados los modelos con nuevos datos.

- Severidad (S^*): 5. Disminuye al anticipar escenarios adversos.
- Probabilidad de ocurrencia (O^*): 4. Se reduce mediante ajustes periódicos de los modelos.

Riesgo 6: fallos técnicos en ejecución de modelos (capacidad limitada en entorno local)

Plan de mitigación específico: preparar entorno de trabajo en la nube como alternativa, reducir el volumen de datos en pruebas intermedias y optimizar código.

- Severidad (S^*): 7. Baja al contar con opciones escalables y respaldo técnico.
- Probabilidad de ocurrencia (O^*): 2. Se reduce con planificación anticipada de infraestructura.

Riesgo 7: dificultad para integrar márgenes y demanda en simulación en el tiempo proyectado

Plan de mitigación específico: modularizar la simulación por componentes (demanda y márgenes por separado), establecer entregables parciales y validar lógica por fases.

- Severidad (S^*): 5. Se reduce por una mejor organización del desarrollo.
- Probabilidad de ocurrencia (O^*): 3. Se mitiga al aplicar integración progresiva de componentes.

14. Sprint Review

La revisión de sprint es una instancia clave en la metodología ágil adoptada, ya que permite validar el cumplimiento parcial del backlog y anticipar posibles desvíos en el desarrollo. A continuación, se detallan las funcionalidades más relevantes de cinco historias de usuario seleccionadas del backlog. Para cada una se identifican las tareas técnicas asociadas, el entregable esperado, los criterios de aceptación, y los principales riesgos u observaciones.

Cuadro 18. Evaluación proyectada de funcionalidades clave del backlog.

HU seleccionada	Tareas asociadas	Entregable esperado	¿Cómo sabrás que está cumplida?	Observaciones o riesgos
HU2	Tarea 1 a 4	Set validado de predictores y datos externos	Variables alineadas con definición y sin problemas de carga	Requiere disponibilidad de fuentes externas
HU2	Tarea 5 a 9	Modelos XGBoost y RF entrenados	Métricas de evaluación aceptables en conjunto de test	Posible sobreajuste o sesgo en los datos
HU3	Tarea 1 a 5	Módulo de márgenes con datos reales	Cálculo reproducible de márgenes unitarios y totales	Validación con equipo financiero es crítica
HU4	Tarea 1 a 6	Simulador de escenarios comerciales	Permite simular cambios de precios y volúmenes con impacto visual	Riesgo de inconsistencias en lógica y validación
HU5	Tarea 1 a 5	Módulo de integración de simulación con predicciones	Panel funcional y verificado con usuarios	Riesgo técnico en compatibilidad de outputs entre modelos
HU9	Tarea 1 a 5	Dashboard principal integrado	Visualización funcional de métricas y filtros clave	Feedback del usuario puede requerir rediseños

15. Sprint Retrospective

La retrospectiva de sprint permite reflexionar y establecer mejoras sobre la dinámica de trabajo durante el desarrollo del proyecto. A partir de esta práctica, se identifican acciones para reforzar lo que funcionó, mejorar lo que no, y planificar nuevas formas de trabajo.

Se presenta a continuación una tabla basada en la herramienta “Estrella de la Retrospectiva”, con foco en tres sprints técnicos clave (1, 2 y 4) y uno no técnico (15), seleccionados por su importancia en la etapa inicial, de entrenamiento de modelos y de cierre del proyecto.

Cuadro 19. Reflexión por Sprint – Estrella de la Retrospectiva.

Sprint	¿Qué hacer más?	¿Qué hacer menos?	¿Qué mantener?	¿Qué empezar a hacer?	¿Qué dejar de hacer?
Técnico – 1	Validaciones frecuentes del modelo con pruebas intermedias	Correcciones sin documentar	Visualización clara de patrones temporales	Registrar cambios en estructura de variables	Iterar sin guardar versiones intermedias
Técnico – 2	Comparaciones explícitas entre modelos con métricas claras	Repeticiones sin control de versiones	Curvas de validación temporal	Documentar supuestos del modelo	Entrenamientos sin control de hiperparámetros
Técnico – 4	Testeo de modelos con inputs alternativos	Ajustes de parámetros sin justificación	Matriz comparativa de resultados entre modelos	Anotar configuraciones óptimas	Cambiar parámetros sin guardar configuración
No técnico – 15	Consolidación visual de la memoria	Cambios innecesarios en versiones finales	Estructura por bloques temáticos	Separar visuales en anexos	Forzar gráficos que no aportan claridad