



Predicción de eventos climáticos y meteorológicos extremos utilizando inteligencia artificial

Autor:

Ing. Sevann Radhak Triztan

Director:

Título y Nombre del director (pertenencia)

Codirector:

Título y Nombre del codirector (FIUBA)

Esta planificación fue realizada en el curso de Gestión de proyectos entre el 18 de junio de 2024 y el 13 de agosto de 2024.

Índice

1. Descripción técnica-conceptual del proyecto a realizar	5
2. Identificación y análisis de los interesados	6
3. Propósito del proyecto	6
4. Alcance del proyecto	6
5. Supuestos del proyecto.	7
6. Requerimientos	8
7. Historias de usuarios (<i>Product backlog</i>).	8
8. Entregables principales del proyecto	9
9. Desglose del trabajo en tareas	10
10. Diagrama de Activity On Node.	11
11. Diagrama de Gantt	13
12. Presupuesto detallado del proyecto	15
13. Gestión de riesgos	15
14. Gestión de la calidad	16
15. Procesos de cierre	17

Registros de cambios

Revisión	Detalles de los cambios realizados	Fecha
0	Creación del documento	18 de junio de 2024
1	Se completa hasta el punto 5 inclusive	2 de julio de 2024
2	Se completa hasta el punto 9 inclusive	9 de julio de 2024
3	Se completa hasta el punto 12 inclusive	30 de julio de 2024

Acta de constitución del proyecto

Buenos Aires, 18 de junio de 2024

Por medio de la presente se acuerda con el Ing. Sevann Radhak Triztan que su Trabajo Final de la Carrera de Especialización en Inteligencia Artificial se titulará “Predicción de eventos climáticos y meteorológicos extremos utilizando inteligencia artificial” y consistirá esencialmente en la implementación de técnicas de inteligencia artificial (IA) para mejorar las predicciones de eventos climáticos y meteorológicos extremos. El trabajo tendrá un presupuesto preliminar estimado de 600 horas y un costo estimado de \$ XXX, con fecha de inicio el 18 de junio de 2024 y fecha de presentación pública el en el mes de diciembre de 2024.

Se adjunta a esta acta la planificación inicial.

Dr. Ing. Ariel Lutenberg
Director posgrado FIUBA

Carlos Roberto Salinas
Dirección de Meteorología e Hidrología
Cliente

Título y Nombre del director
Director del Trabajo Final

1. Descripción técnica-conceptual del proyecto a realizar

El cambio climático y los eventos meteorológicos extremos presentan desafíos críticos a nivel global y local, afectando la vida, la infraestructura y la economía. En Paraguay, los eventos extremos como olas de calor, sequías y precipitaciones intensas han aumentado, destacando la necesidad urgente de mejorar la precisión y la efectividad de los pronósticos meteorológicos.

Las actuales técnicas de monitoreo y predicción no siempre resultan suficientes para emitir advertencias tempranas y precisas. Este proyecto enfrenta el reto de adquirir, analizar y aplicar conocimientos para resolver problemas climatológicos complejos y busca utilizar la inteligencia artificial (IA) para analizar grandes volúmenes de datos meteorológicos y climáticos, tanto actuales como históricos, con el fin de mejorar la precisión de las predicciones y la toma de decisiones.

Este proyecto, propuesto por la Dirección de Meteorología e Hidrología (DINAC) y el Parque Tecnológico Itaipu de Paraguay, busca abordar los desafíos actuales en la predicción de eventos climáticos extremos mediante el uso de herramientas de IA, contribuyendo así a mitigar los impactos del cambio climático en Paraguay y la región.

En la figura 1 se presenta el diagrama en bloques del sistema. Se observan los pasos propuestos para el desarrollo del presente proyecto.

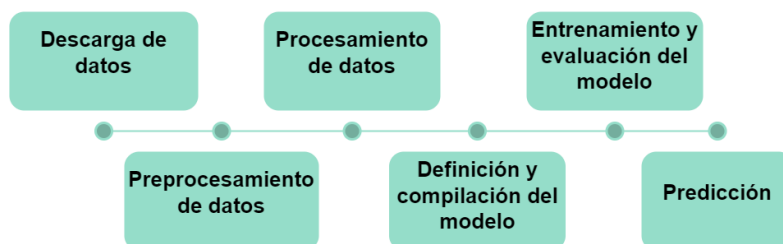


Figura 1. Diagrama en bloques del sistema.

Descarga de datos: la Dirección de Meteorología e Hidrología pondrá a disposición una base de datos para la obtención de datos históricos y actuales de las condiciones meteorológicas.

Preprocesamiento de datos: se llevará a cabo un análisis de datos en crudo con el objetivo de dar tratamiento a datos vacíos, inconsistentes o conflictivos.

Procesamiento de datos: serán aplicadas técnicas como la imputación de valores faltantes, reducción de dimensionalidad, estimación e interpretación de valores representativos, entre otras.

Definición y compilación del modelo: diferentes modelos y técnicas serán propuestas, entrenadas y evaluadas para determinar el mejor modelo a utilizar.

Entrenamiento y evaluación del modelo: el modelo seleccionado será entrenado con los datos procesados y evaluado aplicando técnicas avanzadas y métricas específicas para determinar su precisión y robustez.

Predicción: una vez evaluado y ajustado el modelo, se procederá a realizar predicciones sobre los datos meteorológicos futuros, proporcionando información crítica para la toma de decisiones en la gestión de eventos climáticos extremos.

2. Identificación y análisis de los interesados

Rol	Nombre y Apellido	Organización	Puesto
Cliente	Carlos Roberto Salinas	Dirección de Meteorología e Hidrología	Gerente de Climatología
Product Owner	Carlos Roberto Salinas	Dirección de Meteorología e Hidrología	Gerente de Climatología
Responsable	Ing. Sevann Radhak Triztan	FIUBA	Alumno
Orientador	Título y Nombre del director	pertenencia	Director del Trabajo Final
Usuario final	Dr. Sidney da Silva Viana	Central Hidroeléctrica Itaipu	Gerente de tecnologías en SI

- Cliente: el Dr. Ing. Sidney da Silva Vianna es experto en la temática, ejerce como Gerente de tecnologías en Sistemas de la Información (SI) y va a ayudar con la definición de los requerimientos y el desarrollo de la solución.
- Usuario final: Carlos Roberto Salinas, ejerce como Gerente de Climatología en la Dirección de Meteorología e Hidrología y la Dirección Nacional de Aeronáutica Civil. Proporcionará información útil respecto a las expectativas reales del usuario.

3. Propósito del proyecto

Desarrollar un sistema avanzado de predicción de eventos climáticos y meteorológicos extremos utilizando técnicas de IA. Este sistema tiene como objetivo principal mejorar la precisión y la rapidez de los pronósticos meteorológicos, facilitando una respuesta más efectiva ante los fenómenos climáticos adversos y contribuyendo a la mitigación de sus impactos en Paraguay y sus alrededores.

4. Alcance del proyecto

El proyecto incluye:

- Recolección y preprocesamiento de datos:
 - Obtención de datos meteorológicos históricos y actuales almacenados en la base de datos de la DINAC.
 - Filtrado, normalización y preprocesamiento de los datos para su uso en los modelos de IA.
- Investigación y desarrollo de algoritmos de IA:
 - Identificación y selección de algoritmos de IA adecuados para el análisis de datos meteorológicos.
 - Desarrollo y ajuste de modelos de IA para predicción climática a corto y largo plazo.

- Implementación de la plataforma de predicción:
 - Diseño y desarrollo de una plataforma tecnológica que integre los modelos de IA desarrollados.
 - Implementación de interfaces de usuario para la visualización de predicciones y alertas.
- Validación y optimización de modelos:
 - Pruebas y validación de los modelos de predicción utilizando datos históricos y actuales.
 - Optimización continua de los modelos para mejorar su precisión y eficiencia.
- Gestión y monitoreo del proyecto:
 - Establecimiento de un cronograma detallado y seguimiento del progreso del proyecto.
 - Gestión de riesgos y resolución de problemas durante el desarrollo del proyecto.

El proyecto no incluye:

- Implementación de infraestructura física:
 - La construcción, instalación o conexión a estaciones meteorológicas no está incluida.
 - La actualización o mejora de la infraestructura física existente de recopilación de datos meteorológicos no forma parte del alcance del proyecto.
- Mantenimiento a largo plazo:
 - El mantenimiento continuo de la plataforma y los sistemas desarrollados, después de la finalización del proyecto, no está incluido.
 - Las tareas de mantenimiento o soporte post-implementación.
- Distribución y gestión de alertas:
 - La implementación de sistemas de distribución masiva de alertas no está incluida.
- Integración con sistemas externos:
 - La integración completa con otros sistemas de gestión de emergencias o bases de datos externas fuera del ámbito meteorológico no se incluye.
 - La integración directa con sistemas externos.

5. Supuestos del proyecto

Para el desarrollo del presente proyecto se supone que:

- Se tendrá acceso continuo a una fuente de datos provista por la DINAC.
- La DINAC cuenta con una infraestructura tecnológica adecuada que permitirá el correcto entrenamiento y despliegue de la solución.

- Se fijarán y desarrollarán reuniones periódicas con el cliente para abordar temas relacionados con el proyecto.
- Se mantendrá la relación mediante el programa de vinculación con la DINAC hasta finalizar el desarrollo del proyecto.
- Se contará con la colaboración de expertos: se asume que será posible obtener la colaboración de expertos en el área que se requiera, desde las cuestiones climatológicas hasta las que involucren técnicas de inteligencia artificial.

6. Requerimientos

1. Requerimientos funcionales:

- 1.1. El sistema debe ser capaz de recolectar datos meteorológicos históricos y actuales de la base de datos de la DINAC.
- 1.2. El sistema debe implementar y entrenar múltiples modelos de IA para determinar su conveniencia en términos de precisión, eficiencia y robustez.
- 1.3. El sistema debe superar a los modelos actuales en términos de precisión.
- 1.4. El sistema debe emitir advertencias tempranas basadas en las predicciones de eventos extremos.

2. Requerimientos no funcionales:

- 2.1. La solución debe estar programada de forma modular para que el código de sus características funcionales pueda ser fácilmente reutilizado y/o escalado.
- 2.2. El sistema debe ser capaz de procesar grandes volúmenes de datos sin degradar su rendimiento.

3. Requerimientos de datos:

- 3.1. Los datos meteorológicos históricos y actuales deben estar disponibles en la base de datos de la DINAC.
- 3.2. Deben definirse formatos estándar para la importación y exportación de datos.

4. Requerimientos de documentación:

- 4.1. Debe crearse una documentación técnica del sistema, incluyendo su arquitectura, diseño y funcionamiento.

5. Requerimientos de testing:

- 5.1. La efectividad de los modelos será acordada y evaluada en conjunto con el cliente.

7. Historias de usuarios (*Product backlog*)

Las historias están valoradas en un sistema de puntaje (*story points*) basado en la estimación de tres categorías: dificultad, complejidad e incertidumbre del trabajo en cada una. La escala de *story points* es:

- Trivial: 1

- Bajo: 2
- Medio: 3
- Alto: 5 o más

1. Como Product Owner, quiero obtener datos meteorológicos históricos y actuales de la base de datos de la DINAC para realizar análisis precisos.

Story points: 8 (complejidad: 1, dificultad: 2, incertidumbre: 5)

2. Como Product Owner, quiero limpiar y normalizar los datos meteorológicos para garantizar su calidad ante los modelos de IA que serán entrenados.

Story points: 13 (complejidad: 5, dificultad: 5, incertidumbre: 3)

3. Como Product Owner, quiero aplicar técnicas de imputación de valores faltantes para manejar datos incompletos.

Story points: 8 (complejidad: 3, dificultad: 3, incertidumbre: 2)

4. Como Product Owner, quiero utilizar modelos de IA para predecir eventos climáticos extremos y mejorar la precisión de las predicciones.

Story points: 13 (complejidad: 5, dificultad: 5, incertidumbre: 3)

5. Como Usuario, quiero comparar los resultados obtenidos de los modelos de IA entrenados con el modelo actual para seleccionar el más adecuado para nuestras necesidades.

Story points: 13 (complejidad: 5, dificultad: 3, incertidumbre: 5)

6. Como Product Owner, quiero emitir alertas tempranas de eventos climáticos extremos para tomar decisiones informadas y mitigar impactos.

Story points: 13 (complejidad: 3, dificultad: 5, incertidumbre: 5)

7. Como Product Owner, quiero elaborar una publicación científica basada en los resultados obtenidos del proyecto para contribuir al conocimiento en el campo de la predicción de eventos climáticos extremos.

Story points: 8 (complejidad: 2, dificultad: 3, incertidumbre: 3)

8. Entregables principales del proyecto

- Documento de especificación de requerimientos.
- Modelos de IA entrenados y evaluados.
- Sistema de predicción implementado.
- Informe final del proyecto.
- Código fuente y recursos digitales.
- Memoria del trabajo final.
- Publicación científica basada en los resultados del proyecto.

9. Desglose del trabajo en tareas

1. Grupo de tareas 1: recolección y preprocesamiento de datos (100 h)
 - 1.1. Tarea 1: definir y establecer la conexión con la base de datos de la DINAC (20 h).
 - 1.2. Tarea 2: desarrollar scripts para la extracción automática de datos históricos y actuales (10 h).
 - 1.3. Tarea 3: validar la integridad y calidad de los datos recolectados (10 h).
 - 1.4. Tarea 4: limpieza y normalización de datos para eliminar valores atípicos y errores (25 h).
 - 1.5. Tarea 5: aplicación de técnicas de imputación para manejar datos faltantes (30 h).
 - 1.6. Tarea 6: documentar el proceso de obtención de datos (5 h)
2. Grupo de tareas 2: desarrollo y evaluación de modelos de IA (180 h)
 - 2.1. Tarea 7: investigar y seleccionar algoritmos apropiados para el análisis de datos meteorológicos (30 h).
 - 2.2. Tarea 8: implementar los modelos seleccionados en el entorno de desarrollo (50 h).
 - 1) Subtarea 8.1: configurar el entorno de desarrollo para la implementación de modelos (10 h).
 - 2) Subtarea 8.2: desarrollar el código base para el primer modelo de IA (20 h).
 - 3) Subtarea 8.3: desarrollar el código base para el segundo modelo de IA (20 h).
 - 2.3. Tarea 9: entrenar y ajustar los modelos con datos preprocesados (60 h).
 - 1) Subtarea 9.1: entrenar el primer modelo de IA con los datos preprocesados (20 h).
 - 2) Subtarea 9.2: entrenar el segundo modelo de IA con los datos preprocesados (20 h).
 - 3) Subtarea 9.3: realizar ajustes y optimizaciones en el primer modelo (10 h).
 - 4) Subtarea 9.4: realizar ajustes y optimizaciones en el segundo modelo (10 h).
 - 2.4. Tarea 10: comparar la precisión y eficiencia de los modelos implementados con el modelo actual (40 h).
 - 1) Subtarea 10.1: definir métricas y criterios de evaluación para los modelos (10 h).
 - 2) Subtarea 10.2: realizar pruebas de precisión y eficiencia en el primer modelo (15 h).
 - 3) Subtarea 10.3: realizar pruebas de precisión y eficiencia en el segundo modelo (15 h).
3. Grupo de tareas 3: implementación del sistema y documentación (150 h).
 - 3.1. Tarea 11: diseñar la arquitectura del sistema de predicción de eventos climáticos (40 h).
 - 3.2. Tarea 12: implementar el sistema de predicción de eventos climáticos (40 h).
 - 3.3. Tarea 13: implementar el proceso de emisión de predicciones y alertas (30 h).
 - 3.4. Tarea 14: realizar pruebas de aceptación con el cliente para validar la funcionalidad del sistema (20 h).
 - 3.5. Tarea 15: elaborar la documentación técnica del sistema y los manuales de usuario (20 h).

4. Grupo de tareas 4: gestión del proyecto y presentación (50 h).
 - 4.1. Tarea 16: gestionar riesgos y problemas durante el desarrollo del proyecto (30 h).
 - 4.2. Tarea 17: preparar la presentación pública del proyecto y demostración ante el jurado (20 h).
5. Grupo de tareas 5: elaboración de una publicación científica (60 h).
 - 5.1. Tarea 24: investigación y recopilación de información relevante (10 h).
 - 5.2. Tarea 25: redacción del artículo científico (30 h).
 - 5.3. Tarea 26: revisión y corrección del artículo científico (20 h).
6. Grupo de tareas 6: generación de entregables y proceso de cierre (60 h).
 - 6.1. Tarea 18: inicio de la elaboración de la memoria técnica (20 h).
 - 6.2. Tarea 19: revisión y correcciones de la memoria técnica (5 h).
 - 6.3. Tarea 20: finalización de la elaboración de la memoria técnica (20 h).
 - 6.4. Tarea 21: revisión y correcciones de la memoria técnica final (5 h).
 - 6.5. Tarea 22: elaboración de la presentación final (8 h).
 - 6.6. Tarea 23: revisión y correcciones de la presentación final (2 h).

Cantidad total de horas: 600 h.

10. Diagrama de Activity On Node

En la figura 2 se observa el diagrama de *Activity on Node* del proyecto.

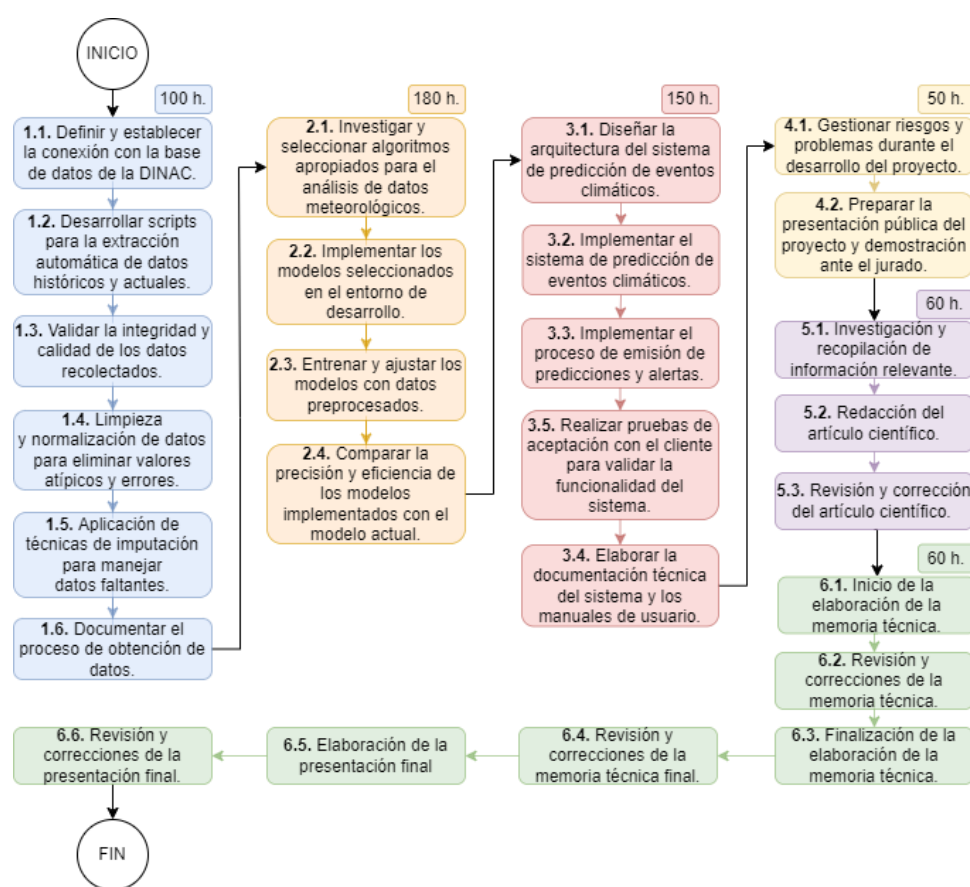


Figura 2. Diagrama de *Activity on Node*.

11. Diagrama de Gantt

En la figura 3 se observa el diagrama de *Gantt* del proyecto.

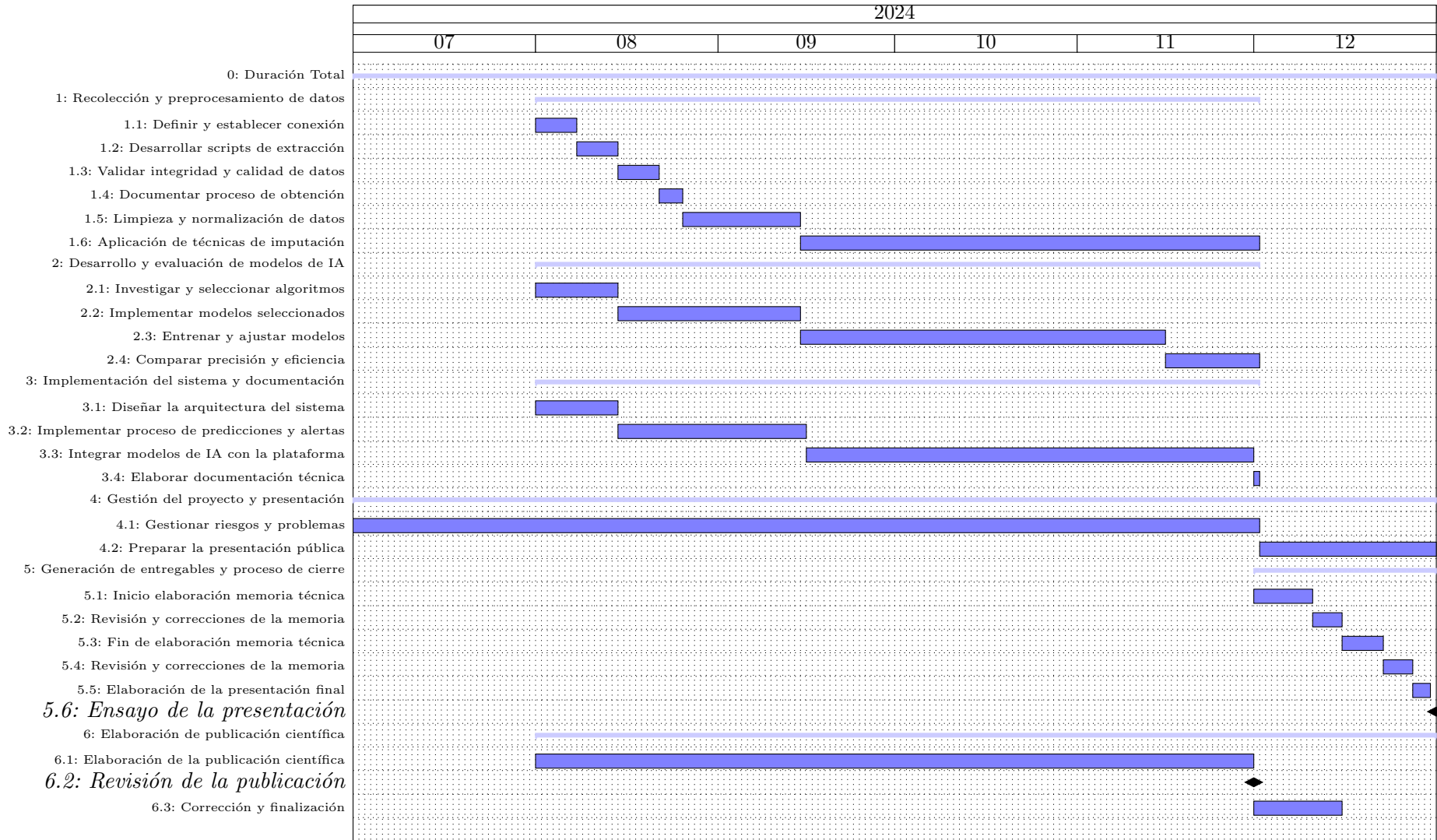


Figura 3. Diagrama de Gantt

12. Presupuesto detallado del proyecto

COSTOS DIRECTOS			
Descripción	Cantidad	Valor unitario	Valor total
Hora de investigación y desarrollo	600 h	\$20 USD	\$12000 USD
Google Colaboratory	330 h	\$0 USD	\$0 USD
SUBTOTAL			\$12000 USD
COSTOS INDIRECTOS			
Descripción	Cantidad	Valor unitario	Valor total
Notebook	1 unidad	\$1200 USD	\$1200 USD
Imprevistos y contingencias (30% de costos directos)	-	-	\$3600 USD
SUBTOTAL			\$4800 USD
TOTAL			\$16800 USD

13. Gestión de riesgos

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

Riesgo 1: detallar el riesgo (riesgo es algo que si ocurre altera los planes previstos de forma negativa)

- Severidad (S): mientras más severo, más alto es el número (usar números del 1 al 10). Justificar el motivo por el cual se asigna determinado número de severidad (S).
- Probabilidad de ocurrencia (O): mientras más probable, más alto es el número (usar del 1 al 10). Justificar el motivo por el cual se asigna determinado número de (O).

Riesgo 2:

- Severidad (S): X.
Justificación...
- Ocurrencia (O): Y.
Justificación...

Riesgo 3:

- Severidad (S): X.
Justificación...
- Ocurrencia (O): Y.
Justificación...

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

Riesgo	S	O	RPN	S*	O*	RPN*

Criterio adoptado:

Se tomarán medidas de mitigación en los riesgos cuyos números de RPN sean mayores a...

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:

Riesgo 1: plan de mitigación (si por el RPN fuera necesario elaborar un plan de mitigación).
Nueva asignación de S y O, con su respectiva justificación:

- Severidad (S*): mientras más severo, más alto es el número (usar números del 1 al 10). Justificar el motivo por el cual se asigna determinado número de severidad (S).
- Probabilidad de ocurrencia (O*): mientras más probable, más alto es el número (usar del 1 al 10). Justificar el motivo por el cual se asigna determinado número de (O).

Riesgo 2: plan de mitigación (si por el RPN fuera necesario elaborar un plan de mitigación).

Riesgo 3: plan de mitigación (si por el RPN fuera necesario elaborar un plan de mitigación).

14. Gestión de la calidad

Elija al menos diez requerimientos que a su criterio sean los más importantes/críticos/que aportan más valor y para cada uno de ellos indique las acciones de verificación y validación que permitan asegurar su cumplimiento.

- Req #1: copiar acá el requerimiento con su correspondiente número.
 - Verificación para confirmar si se cumplió con lo requerido antes de mostrar el sistema al cliente. Detallar.
 - Validación con el cliente para confirmar que está de acuerdo en que se cumplió con lo requerido. Detallar.

Tener en cuenta que en este contexto se pueden mencionar simulaciones, cálculos, revisión de hojas de datos, consulta con expertos, mediciones, etc.

Las acciones de verificación suelen considerar al entregable como “caja blanca”, es decir se conoce en profundidad su funcionamiento interno.

En cambio, las acciones de validación suelen considerar al entregable como “caja negra”, es decir, que no se conocen los detalles de su funcionamiento interno.

15. Procesos de cierre

Establecer las pautas de trabajo para realizar una reunión final de evaluación del proyecto, tal que contemple las siguientes actividades:

- Pautas de trabajo que se seguirán para analizar si se respetó el Plan de Proyecto original:
 - Indicar quién se ocupará de hacer esto y cuál será el procedimiento a aplicar.
- Identificación de las técnicas y procedimientos útiles e inútiles que se emplearon, los problemas que surgieron y cómo se solucionaron:
 - Indicar quién se ocupará de hacer esto y cuál será el procedimiento para dejar registro.
- Indicar quién organizará el acto de agradecimiento a todos los interesados, y en especial al equipo de trabajo y colaboradores:
 - Indicar esto y quién financiará los gastos correspondientes.