

Algoritmo Predictivo de Enfermedades Fundadas en Cultivos de Café Utilizando Aprendizaje Profundo y Sensores de Campo

**CONVOCATORIA COLOMBIA INTELIGENTE: CIENCIA Y TECNOLOGÍAS
CUÁNTICAS E INTELIGENCIA ARTIFICIAL PARA LOS TERRITORIOS |
Convocatoria 966**

COTECMAR - December 2025

Tabla de Contenido

1. Generalidades del Proyecto	3
2. Resumen Ejecutivo	3
3. Planteamiento del Problema y Justificación	4
4. Marco Teórico y Estado del Arte	4
4.1. <i>Introducción al Dominio</i>	4
4.2. <i>Revisión de la Literatura (Literature Review)</i>	5
4.3. <i>Tecnologías y Enfoques Actuales (State of the Art)</i>	6
4.4. <i>Brechas de Conocimiento y Oportunidades (Knowledge Gaps & Opportunities)</i>	6
5. Objetivos	7
6. Metodología Propuesta	8
7. Plan de Ejecución y Gestión	9
7.1. <i>Cronograma de Actividades</i>	9
7.2. <i>Matriz de Riesgos</i>	11
8. Resultados e Impactos Esperados	12
8.1. <i>Resultados Esperados (Entregables)</i>	12
8.2. <i>Impactos Esperados</i>	12
9. Referencias Bibliográficas	13

1. Generalidades del Proyecto

Título: Algoritmo Predictivo de Enfermedades Fúngicas en Cultivos de Café Utilizando Aprendizaje Profundo y Sensores de Campo

Convocatoria: CONVOCATORIA COLOMBIA INTELIGENTE: CIENCIA Y TECNOLOGÍAS CUÁNTICAS E INTELIGENCIA ARTIFICIAL PARA LOS TERRITORIOS | Convocatoria 966

Programa: Por definir

Entidad/Persona: COTECMAR

Línea Temática: Tecnologías Cuánticas, Inteligencia Artificial, Territorios, CTel, Ciencia

Duración: Por definir

Área OCDE: Por definir

● **Descripción:** Las enfermedades fúngicas, como la Roya del cafeto, causan pérdidas significativas y afectan la subsistencia de miles de pequeños productores de café en Colombia. Este proyecto propone desarrollar un algoritmo de aprendizaje profundo, alimentado con datos de sensores IoT climáticos y espectrales en campo, para la detección temprana y predicción del riesgo de brotes de enfermedades fúngicas. Se realizará un piloto en fincas cafeteras de Cundinamarca y Santander para validar la eficacia del sistema en condiciones reales y calibrar los modelos predictivos. El objetivo es reducir las pérdidas por enfermedades en al menos un 25% y optimizar el uso de fungicidas, contribuyendo a la sostenibilidad de la caficultura.

● **Palabras Clave:** Inteligencia Artificial, Aprendizaje Profundo, Sensores IoT, Caficultura, Enfermedades Fúngicas, Agricultura de Precisión, Territorios

2. Resumen Ejecutivo

La agricultura moderna se enfrenta a la amenaza constante de enfermedades fúngicas, como la Roya del cafeto, que causan pérdidas económicas multimillonarias y comprometen la subsistencia de miles de pequeños productores en Colombia. La gestión actual es reactiva e inefficiente. Nuestro proyecto propone una solución innovadora: desarrollar y validar un algoritmo predictivo robusto que, utilizando aprendizaje profundo y sensores de campo, detectará tempranamente el riesgo de brotes de enfermedades fúngicas en cultivos de café, optimizando su gestión y reduciendo drásticamente las pérdidas económicas.

Para lograrlo, diseñaremos e implementaremos un sistema IoT para la recolección automatizada de datos climáticos y espectrales de alta calidad. Posteriormente, desarrollaremos y entrenaremos un modelo de aprendizaje profundo, como redes neuronales convolucionales o recurrentes, para predecir el riesgo de brotes con una precisión superior al 85% y con semanas de anticipación. Finalmente, validaremos y optimizaremos este algoritmo en condiciones reales de campo, buscando una tasa de acierto del 90% antes

de la aparición visible de síntomas. Todo esto se guiará por la metodología estructurada CRISP-DM.

Los entregables clave incluyen un sistema IoT funcional, un algoritmo predictivo entrenado y validado en laboratorio, y una versión optimizada para campo. El impacto técnico-científico avanzará la agricultura de precisión en Colombia, generando nuevo conocimiento sobre los patrones de la Roya. Económicamente, los caficultores reducirán costos por fungicidas, disminuirán pérdidas de cosecha y aumentarán su rentabilidad. Social y ambientalmente, mejoraremos la calidad de vida de las comunidades y fomentaremos una caficultura más sostenible al reducir la carga química en los ecosistemas.

Este proyecto representa una inversión estratégica fundamental para la resiliencia y competitividad del sector caficultor colombiano, transformando la gestión de enfermedades de reactiva a proactiva y sentando un precedente para la agricultura del futuro.

3. Planteamiento del Problema y Justificación

La agricultura moderna se enfrenta a desafíos significativos, incluyendo la creciente demanda de alimentos y la amenaza constante de plagas y enfermedades. En este contexto, las enfermedades fúngicas representan una de las principales causas de pérdidas de rendimiento en cultivos a nivel global, afectando la seguridad alimentaria y la economía de los agricultores. Específicamente, en la caficultura colombiana, enfermedades como la Roya del cafeto (*Hemileia vastatrix*) pueden devastar plantaciones enteras, impactando directamente la subsistencia de miles de pequeños productores y generando pérdidas económicas multimillonarias que comprometen la viabilidad a largo plazo de este sector estratégico. La gestión actual, a menudo reactiva, resulta en un uso ineficiente de fungicidas y una mitigación tardía de los brotes.

Aunque la literatura científica (Chatla Subbarayudu & Mohan Kubendiran, 2024; Amudha S et al., 2023; Engr. Faiza Irfan et al., 2025) demuestra avances significativos en el uso del aprendizaje profundo y los sensores IoT para la detección de enfermedades en una variedad de cultivos, y trabajos como los de Deepak Thakur et al. (2024) y Raveena S y S. R (2023) abordan la Roya del cafeto, existe una brecha crítica. La mayoría de los enfoques se centran en la detección de síntomas ya presentes, y no en la predicción temprana del riesgo de brote utilizando una combinación de datos climáticos y espectrales antes de la aparición visual de la enfermedad. Además, la aplicación específica y la validación de sistemas predictivos robustos para enfermedades fúngicas en cultivos de café bajo las condiciones agroecológicas particulares de regiones como Cundinamarca y Santander en Colombia aún presentan un campo fértil para la investigación y el desarrollo, lo que limita la capacidad de los agricultores para implementar intervenciones proactivas y eficientes.

Este proyecto,

4. Marco Teórico y Estado del Arte

4.1. Introducción al Dominio

La agricultura moderna se enfrenta a desafíos significativos, incluyendo la creciente demanda de alimentos y la amenaza constante de plagas y enfermedades. Las enfermedades fúngicas, en particular, representan una de las principales causas de pérdidas de rendimiento en cultivos a nivel global, afectando la seguridad alimentaria y la economía de los agricultores (Chatla Subbarayudu & Mohan Kubendir, 2024). En el de las principales causas de pérdidas de rendimiento en cultivos a nivel global, afectando la seguridad alimentaria y la economía de los agricultores (Chatla Subbarayudu & Mohan Kubendir, 2024). En el contexto de la caficultura, enfermedades como la Roya del cafeto (*Hemileia vastatrix*) pueden devastar plantaciones enteras, impactando directamente la subsistencia de miles de pequeños productores. Para contrarrestar estas amenazas, la agricultura de precisión ha emergido como un paradigma transformador, utilizando tecnologías avanzadas para optimizar la gestión de cultivos y recursos.

Dentro de la agricultura de precisión, el aprendizaje profundo (Deep Learning), una rama de la inteligencia artificial, ha demostrado ser particularmente eficaz en tareas complejas como la detección y clasificación de enfermedades vegetales a partir de imágenes y datos ambientales. Paralelamente, el Internet de las Cosas (IoT) ha revolucionado la recolección de datos en campo, permitiendo el monitoreo en tiempo real de variables climáticas, edáficas y de salud de las plantas a través de una red de sensores interconectados (Engr. Faiza Irfan et al., 2025). La sinergia entre el aprendizaje profundo y los sensores IoT ofrece una oportunidad sin precedentes para desarrollar sistemas predictivos que permitan una detección temprana y una intervención proactiva contra las enfermedades fúngicas, mejorando la sostenibilidad y productividad agrícola.

4.2. Revisión de la Literatura (Literature Review)

La investigación en la detección y predicción de enfermedades en cultivos ha avanzado significativamente gracias al aprendizaje profundo y la integración de sensores. **Chatla Subbarayudu y Mohan Kubendir (2024)** realizaron una exhaustiva revisión de las técnicas de Machine Learning y Deep Learning aplicadas a la predicción de enfermedades en cultivos. Su estudio concluye que las técnicas de aprendizaje profundo superan en eficiencia a los enfoques tradicionales de Machine Learning, destacando su potencial para prevenir pérdidas de cosechas al identificar enfermedades en etapas tempranas. Este trabajo subraya la importancia de los modelos de Deep Learning como columna vertebral para sistemas de detección avanzados.

En un contexto similar, **Amudha S et al. (2023)** propusieron un sistema modular para la predicción del rendimiento de cultivos y la detección automática de enfermedades. Utilizando sensores IoT para parámetros como Nitrógeno-Potasio-Fósforo (NPP), humedad, temperatura y pH, combinado con cámaras para el análisis de hojas, su modelo basado en CNN, ResNet, VGGNet y MobileNet logró una precisión del 98.2% en la identificación de enfermedades en 14 especies de plantas. Este estudio es relevante porque demuestra la eficacia de la combinación de datos de sensores ambientales con análisis de imágenes para el diagnóstico de enfermedades, un pilar fundamental del proyecto propuesto.

La integración de sensores IoT para la monitorización de la salud de los cultivos es explorada por **Engr. Faiza Irfan et al. (2025)**, quienes presentaron un marco de agricultura inteligente impulsado por IoT. Su implementación piloto en un cultivo de trigo demostró una reducción del 30% en el uso de agua y una precisión del 92% en la detección temprana de enfermedades. Este trabajo valida la viabilidad y los beneficios de los sistemas IoT en la agricultura de precisión, destacando su capacidad para optimizar recursos y mejorar

la sostenibilidad. De manera complementaria, **Dr. Manjusha Tatiya et al. (2025)** enfatizan cómo la convergencia de sensores inteligentes, IA e IoT es una solución sostenible para optimizar la productividad agrícola, permitiendo la adquisición de datos en tiempo real y análisis predictivos para la gestión de plagas y enfermedades.

Específicamente para la caficultura, **Deepak Thakur et al. (2024)** introdujeron SUNet, un modelo híbrido de aprendizaje profundo para la detección y clasificación de enfermedades foliares del café, incluyendo la Roya. SUNet integra U-Net con el sistema de codificación de Segnet y VGG16 para la extracción de características robustas, logrando mejoras significativas en métricas como precisión y F1-score. Este estudio valida la aplicación de arquitecturas avanzadas de Deep Learning para enfermedades específicas del café. Asimismo, **Raveena S y S. R (2023)** propusieron un modelo basado en Deep Belief Networks (DBN) para la predicción de la Roya del cafeto, logrando alta precisión en la detección de áreas dañadas en hojas de café. Estos trabajos demuestran la viabilidad técnica de usar Deep Learning para abordar la Roya del cafeto.

4.3. Tecnologías y Enfoques Actuales (State of the Art)

El estado del arte en la predicción de enfermedades fúngicas en cultivos se caracteriza por la convergencia de la Inteligencia Artificial, particularmente el Aprendizaje Profundo, y la tecnología IoT. Los algoritmos de aprendizaje profundo, especialmente las Redes Neuronales Convolucionales (CNNs), son el pilar para el análisis de imágenes de cultivos, permitiendo la identificación de patrones sutiles asociados con las primeras etapas de las enfermedades. Modelos como ResNet, VGGNet, MobileNet (Amudha S et al., 2023), U-Net y Mask R-CNN (Deepak Thakur et al., 2024) han demostrado ser altamente efectivos para la clasificación y segmentación de enfermedades foliares. La tendencia actual es hacia arquitecturas híbridas que combinan diferentes modelos para mejorar la robustez y precisión.

En paralelo, los sistemas de sensores IoT son fundamentales para la recolección de datos multimodales en tiempo real. Estos incluyen sensores climáticos (temperatura, humedad, precipitación), sensores de suelo (humedad, pH, nutrientes como NPK) y sensores espectrales que capturan información no visible al ojo humano, revelando cambios fisiológicos en las plantas antes de la manifestación de síntomas visuales. La combinación de imágenes RGB con datos espectrales y ambientales es una práctica emergente que potencia la capacidad predictiva de los modelos de aprendizaje profundo (academic_search, 2023). La infraestructura IoT permite la transmisión de estos datos a plataformas en la nube, donde los algoritmos de IA procesan la información para generar alertas y recomendaciones.

Actualmente, existe un énfasis en el desarrollo de sistemas que no solo detecten la enfermedad, sino que también predigan el riesgo de brotes futuros basándose en las condiciones ambientales y los datos históricos. Esto se logra a menudo mediante el uso de modelos de series temporales integrados con arquitecturas de aprendizaje profundo. La implementación de dispositivos de borde (edge devices) con capacidades de inferencia de IA está ganando terreno para permitir el procesamiento de datos directamente en el campo, reduciendo la latencia y la dependencia de la conectividad constante a la nube (Raka Thoriq Araaf et al., 2024).

4.4. Brechas de Conocimiento y Oportunidades (Knowledge Gaps & Opportunities)

A pesar de los avances significativos en el uso del aprendizaje profundo y los sensores IoT para la detección de enfermedades en cultivos, existen brechas de conocimiento y oportunidades que el proyecto busca abordar. Si bien hay una vasta literatura sobre la detección de enfermedades en una variedad de cultivos, la aplicación específica y la validación de sistemas predictivos de enfermedades fúngicas en **cultivos de café bajo las condiciones agroecológicas particulares de regiones como Cundinamarca y Santander en Colombia** aún presentan un campo fértil para la investigación y el desarrollo. La mayoría de los estudios se centran en la detección basada en imágenes de síntomas ya presentes, en lugar de la **predicción temprana del riesgo de brote** utilizando una combinación de datos climáticos y espectrales antes de la aparición visual de la enfermedad.

Otra brecha importante reside en la **integración holística y la calibración de modelos multimodales** que combinen eficazmente datos de sensores climáticos y espectrales con algoritmos de aprendizaje profundo para generar alertas predictivas robustas. Aunque existen componentes individuales, la creación de un sistema unificado que sea validado en condiciones de campo reales y que demuestre una **reducción cuantificable de pérdidas y una optimización del uso de fungicidas** es una necesidad insatisfecha. El proyecto tiene la oportunidad de establecer un marco integral que no solo mejore la precisión de la predicción, sino que también contribuya directamente a la sostenibilidad de la caficultura al reducir la dependencia de intervenciones reactivas y el uso excesivo de agroquímicos.

5. Objetivos

Objetivo General

Desarrollar y validar un algoritmo predictivo robusto para la detección temprana del riesgo de enfermedades fúngicas en cultivos de café, utilizando aprendizaje profundo y datos de sensores de campo, con el fin de optimizar la gestión y reducir las pérdidas económicas en regiones caficultoras de Colombia.

Objetivos Específicos

1. **Objetivo:** Diseñar e implementar un sistema para la recolección y preprocesamiento automatizado de datos climáticos y espectrales de campo.

● **Específico (S):** El equipo de investigación diseñará y configurará una red de sensores IoT en fincas cafetaleras de Cundinamarca y Santander para recolectar datos de temperatura, humedad, precipitación y reflectancia espectral. Se establecerán protocolos para la limpieza y estructuración de estos datos.

● **Medible (M):** Se logrará una tasa de recolección de datos del 98% de los sensores instalados semanalmente. Los datos preprocesados cumplirán con estándares de calidad definidos (e.g., menos del 2% de valores atípicos o faltantes) para su uso en el modelo de aprendizaje profundo.

● **Alcanzable (A):** Es alcanzable dado el avance en tecnologías de sensores IoT y plataformas de gestión de datos, y la experiencia del equipo en instrumentación de campo.

● **Relevante (R):** La disponibilidad de datos de alta calidad es fundamental para entrenar cualquier modelo de aprendizaje profundo y es el primer paso para la predicción temprana de enfermedades,

abordando la limitación actual de gestión reactiva.

- **Plazo (T):** Este objetivo se completará en los primeros 6 meses del proyecto.

2. **Objetivo:** Desarrollar y entrenar un modelo de aprendizaje profundo capaz de predecir el riesgo de brotes de enfermedades fúngicas en café.

- **Específico (S):** El equipo de científicos de datos diseñará, implementará y entrenará una arquitectura de red neuronal convolucional (CNN) o recurrente (RNN) utilizando los datos climáticos y espectrales preprocesados para identificar patrones predictivos de la Roya del cafeto en las regiones de estudio.

- **Medible (M):** El modelo alcanzará una precisión de predicción superior al 85% para la identificación del riesgo de brote de Roya del cafeto con al menos 2 semanas de anticipación, medido mediante métricas como F1-score y AUC en un conjunto de datos de validación independiente.

- **Alcanzable (A):** Es alcanzable considerando la disponibilidad de librerías de aprendizaje profundo, la capacidad computacional y la base de datos robusta generada en la fase inicial del proyecto.

- **Relevante (R):** Este objetivo es el núcleo del proyecto, ya que el modelo predictivo es la herramienta principal para la detección temprana y proactiva de la enfermedad, abordando directamente la brecha crítica identificada de falta de predicción temprana.

- **Plazo (T):** Este objetivo se completará entre el mes 7 y el mes 15 del proyecto.

3. **Objetivo:** Validar el rendimiento del algoritmo predictivo en condiciones reales de campo y optimizar su desempeño.

- **Específico (S):** El algoritmo predictivo será implementado en un entorno de prueba en al menos 3 fincas cafetaleras seleccionadas en Cundinamarca y Santander. Se compararán las predicciones del modelo con la aparición real de síntomas de Roya del cafeto observados y confirmados por agrónomos de campo.

- **Medible (M):** El modelo demostrará una reducción del 20% en falsos positivos y falsos negativos respecto a la validación inicial del modelo, y se logrará una tasa de acierto del 90% en la predicción del riesgo de brote 10 días antes de la aparición visible de síntomas en las fincas piloto.

- **Alcanzable (A):** La validación en campo es una etapa estándar en proyectos de I+D agrícola y es factible con la colaboración de agricultores y agrónomos locales, así como la infraestructura de sensores ya establecida.

- **Relevante (R):** La validación en campo asegura que el algoritmo sea práctico, robusto y efectivo bajo las condiciones agroecológicas específicas de Colombia, garantizando su aplicabilidad y contribuyendo a la implementación de intervenciones proactivas y eficientes.

- **Plazo (T):** Este objetivo se llevará a cabo entre el mes 16 y el mes 24 del proyecto.

6. Metodología Propuesta

Framework Seleccionado: CRISP-DM (Cross-Industry Standard Process for Data Mining)

El framework CRISP-DM ha sido seleccionado como la metodología principal para este proyecto debido a su idoneidad para iniciativas de investigación y desarrollo centradas en el análisis de datos y el aprendizaje automático. Este enfoque iterativo y estructurado se alinea perfectamente con los objetivos del proyecto, facilitando la transición desde la recolección de datos hasta el despliegue de un modelo predictivo robusto. Específicamente, CRISP-DM permitirá abordar de manera sistemática el **Objetivo 1** al guiar la comprensión y preparación de los datos de sensores; el **Objetivo 2** al proporcionar un marco claro para el modelado y entrenamiento del algoritmo de aprendizaje profundo; y el **Objetivo 3** al asegurar una evaluación rigurosa y un despliegue efectivo del algoritmo en condiciones reales de campo, garantizando así la validez y aplicabilidad de los resultados.

Fases Principales de la Metodología:

- **Fase 1: Comprensión del Negocio (Business Understanding)** - Se definen los objetivos del proyecto desde una perspectiva de negocio y se transforman en problemas de minería de datos, identificando los criterios de éxito y los requisitos específicos para la predicción de enfermedades fúngicas.
- **Fase 2: Comprensión de los Datos (Data Understanding)** - Se realiza una recolección inicial de datos, una exploración exhaustiva para familiarizarse con ellos, se identifican problemas de calidad y se descubren patrones iniciales que puedan ser relevantes para el modelado.
- **Fase 3: Preparación de los Datos (Data Preparation)** - Se llevan a cabo las actividades de limpieza, transformación, integración y formateo de los datos recolectados de los sensores para dejarlos listos y óptimos para la fase de modelado, asegurando su calidad y consistencia.
- **Fase 4: Modelado (Modeling)** - Se seleccionan y aplican diversas técnicas de aprendizaje profundo, como redes neuronales convolucionales o recurrentes, y se ajustan sus parámetros para desarrollar y entrenar el algoritmo predictivo de riesgo de brotes de Roya del cafeto.
- **Fase 5: Evaluación (Evaluation)** - Se evalúa a fondo el rendimiento del modelo entrenado frente a los objetivos del negocio y los criterios de éxito definidos, verificando su precisión y robustez en conjuntos de datos de validación independientes.
- **Fase 6: Despliegue (Deployment)** - Se implementa el algoritmo predictivo validado en un entorno de prueba en fincas cafetaleras reales, se monitorea su desempeño en campo y se realizan los ajustes necesarios para asegurar su operatividad y eficacia a largo plazo.

7. Plan de Ejecución y Gestión

Cronograma de Actividades

7.1. Cronograma de Actividades

Fase	Actividad / Hito Clave	Entregable Principal	Duración Estimada (Semanas)
Fase 1: Comprensión del Negocio	<i>Definición de los objetivos del proyecto y los requisitos de negocio para la predicción de enfermedades fúngicas.</i>		4
	1.1. Definición de Objetivos, Alcance y Requisitos del Proyecto	Documento de Requisitos de Negocio y Objetivos del Proyecto	2
	1.2. Análisis de Viabilidad Técnica y Definición de Criterios de Éxito	Informe de Viabilidad y KPIs de Éxito	2
Fase 2: Comprensión de los Datos	<i>Diseño del sistema de recolección de datos y análisis exploratorio inicial de la información.</i>		8
	2.1. Diseño Detallado de la Red de Sensores IoT y Plan de Recolección	Plan de Diseño de Sensores y Protocolos de Recolección	3
	2.2. Instalación y Configuración Inicial de la Red de Sensores en Campo	Red de Sensores IoT Operativa en Fincas Piloto	3
	2.3. Recolección Inicial de Datos y Análisis Exploratorio	Informe de Análisis Exploratorio de Datos Iniciales	2
Fase 3: Preparación de los Datos	<i>Limpieza, transformación e integración de los datos recolectados para el modelado de aprendizaje profundo.</i>		12
	3.1. Diseño e Implementación de Pipelines de Limpieza y Preprocesamiento	Scripts de Preprocesamiento de Datos V1.0	4
	3.2. Integración y Estructuración de Bases de Datos Climáticos y Espectrales	Base de Datos Integrada y Estructurada	4
	3.3. Ingeniería de Características y Generación de Conjuntos de Datos para Modelado	Conjuntos de Datos Limpios y Listos para Entrenamiento	4
Fase 4: Modelado	<i>Desarrollo y entrenamiento de arquitecturas de aprendizaje profundo para la predicción de riesgo de Roya.</i>		24
	4.1. Selección y Diseño de Arquitecturas de Redes Neuronales (CNN/RNN)	Documento de Diseño de Arquitectura del Modelo	6
	4.2. Implementación y Entrenamiento Inicial del Modelo Predictivo de Roya	Modelo de Aprendizaje Profundo Entrenado V1.0	10
	4.3. Optimización de Hiperparámetros y Refinamiento del Modelo	Modelo Optimizado V1.0 (Predicción >85% precisión)	8
Fase 5: Evaluación	<i>Validación rigurosa del rendimiento del modelo frente a los objetivos de precisión y robustez.</i>		12
	5.1. Validación del Rendimiento del Modelo con Conjuntos de Datos Independientes	Informe de Métricas de Rendimiento (F1-score, AUC)	5
	5.2. Análisis Detallado de Robustez y Generalización del Modelo	Reporte de Robustez y Generalización del Modelo	4
	5.3. Generación de Reporte de Evaluación Final del Modelo	Documento de Evaluación y Criterios de Aceptación del Modelo	3

Fase 6: Despliegue	<i>Implementación del algoritmo en campo, monitoreo de desempeño y optimización continua.</i>		36
	6.1. Implementación Piloto del Algoritmo en Entorno de Prueba en Fincas	Algoritmo Predictivo Operativo en Fincas Piloto	8
	6.2. Monitoreo Continuo del Desempeño en Campo y Recolección de Datos de Validación	Datos de Monitoreo de Campo y Registro de Predicciones	14
	6.3. Optimización del Algoritmo Basada en Datos Reales y Retroalimentación de Agrónomos	Versión Optimizada del Algoritmo (Reducción 20% FP/FN)	10
	6.4. Generación de Reporte Final de Rendimiento y Recomendaciones	Informe Final de Rendimiento y Recomendaciones de Implementación	4

Matriz de Riesgos

7.2. Matriz de Riesgos

#	Riesgo Potencial	Probabilidad	Impacto	Estrategia de Mitigación
1	El modelo de aprendizaje profundo no alcanza la precisión predictiva requerida (>85%) <i>Relacionado con: Fase 4: Modelado (4.3) y Fase 5: Evaluación (5.1)</i>	Medium	High	Implementar un marco de evaluación robusto con métricas claras desde el inicio. Explorar múltiples arquitecturas de modelos (ensemble methods) y técnicas de regularización. Disponer de un equipo de expertos en ML para optimización continua y reentrenamiento con nuevos datos. Considerar modelos alternativos de menor complejidad como plan de contingencia.
2	Fallos en la red de sensores IoT o inconsistencia en la recolección de datos en campo <i>Relacionado con: Fase 2: Comprensión de los Datos (2.2, 2.3) y Fase 6: Despliegue (6.2)</i>	Medium	Medium	Realizar pruebas piloto exhaustivas de los sensores antes del despliegue masivo. Implementar redundancia en puntos críticos. Desarrollar un sistema de monitoreo remoto de la salud de los sensores y protocolos de mantenimiento preventivo y correctivo rápido. Capacitar al personal de campo para la verificación y resolución de problemas básicos.
3	Retrasos significativos en la integración y preparación de datos, impactando las fases posteriores de modelado <i>Relacionado con: Fase 3: Preparación de los Datos (3.1, 3.2, 3.3)</i>	Medium	Medium	Establecer hitos intermedios con entregables claros y revisiones periódicas. Asignar recursos dedicados y experimentados en ingeniería de datos. Utilizar herramientas de automatización para la limpieza y preprocesamiento de datos. Mantener un buffer de tiempo en el cronograma para estas actividades complejas.

4	Dificultades o baja aceptación en la implementación piloto del algoritmo en fincas Relacionado con: Fase 6: Despliegue (6.1, 6.3)	Medium	High	Involucrar a los agrónomos y agricultores desde las fases iniciales del diseño. Realizar talleres de capacitación y demostraciones del valor del algoritmo. Recopilar activamente retroalimentación de los usuarios y priorizar ajustes en el algoritmo y la interfaz. Desplegar de forma gradual, comenzando con fincas colaboradoras.
5	Insuficiencia o baja calidad de los datos de etiquetado para el entrenamiento efectivo del modelo Relacionado con: Fase 3: Preparación de los Datos (3.3) y Fase 4: Modelado (4.2)	Medium	High	Establecer protocolos estrictos de etiquetado y validación de datos con expertos agrónomos. Implementar técnicas de aumento de datos (data augmentation). Explorar el aprendizaje semi-supervisado o el uso de modelos pre-entrenados con transferencia de conocimiento si la cantidad de datos etiquetados es limitada. Planificar campañas de recolección de datos adicionales si es necesario.

8. Resultados e Impactos Esperados

8.1. Resultados Esperados (Entregables)

- **Sistema Integrado de Recolección y Preprocesamiento de Datos de Campo:** Un sistema IoT funcional compuesto por una red de sensores instalados en fincas cafetaleras para la adquisición automatizada de datos climáticos y espectrales, junto con una base de datos estructurada que contenga los datos preprocesados y validados, listos para el entrenamiento del modelo. Este entregable corresponde al Objetivo Específico 1.
- **Algoritmo Predictivo de Aprendizaje Profundo Entrenado y Validado en Laboratorio:** Un modelo de aprendizaje profundo (red neuronal convolucional o recurrente) desarrollado, entrenado y validado internamente con una precisión superior al 85% para la predicción temprana del riesgo de brotes de la Roya del cafeto. Este entregable corresponde al Objetivo Específico 2.
- **Algoritmo Predictivo Optimizado y Protocolos de Implementación en Campo:** La versión final del algoritmo predictivo, optimizada tras la validación en condiciones reales de campo, demostrando una reducción del 20% en falsos positivos/negativos y una tasa de acierto del 90%. Incluirá un informe técnico detallado de la validación en fincas piloto y directrices operativas para su implementación y uso por parte de los agricultores y agrónomos. Este entregable corresponde al Objetivo Específico 3.

8.2. Impactos Esperados

- **Impacto Técnico/Científico:**

Este proyecto avanzará significativamente el estado del arte en la agricultura de precisión en Colombia, al establecer una metodología robusta para la integración de datos de sensores IoT con técnicas avanzadas de

aprendizaje profundo para la detección temprana y proactiva de enfermedades fúngicas. Se generará nuevo conocimiento sobre los patrones predictivos de la Roya del cafeto en diversas condiciones agroecológicas, lo que sentará las bases para el desarrollo de futuras herramientas de manejo fitosanitario basadas en datos y extenderá la aplicabilidad de estas tecnologías a otros cultivos y patologías.

● **Impacto Económico:**

El desarrollo y la implementación del algoritmo predictivo tendrán un impacto económico directo y sustancial para los caficultores. Al permitir una detección temprana del riesgo de brote de la Roya con al menos 10 días de anticipación, se optimizará la aplicación de fungicidas, reduciendo su uso indiscriminado y, por ende, los costos asociados. Esto se traducirá en una disminución de las pérdidas de cosecha, un aumento en la productividad y la estabilidad de los rendimientos, mejorando la rentabilidad de las fincas cafetaleras y fortaleciendo la competitividad del café colombiano en los mercados nacionales e internacionales.

● **Impacto Social/Ambiental:**

Socialmente, el proyecto contribuirá a mejorar la calidad de vida y la resiliencia económica de las comunidades caficultoras, especialmente pequeños y medianos productores, al brindarles una herramienta tecnológica avanzada para proteger sus cultivos y medios de subsistencia. Esto fomentará la adopción de prácticas agrícolas más eficientes y la capacitación en nuevas habilidades digitales y de análisis de datos. Desde una perspectiva ambiental, la optimización en el uso de fungicidas reducirá la carga química en los ecosistemas cafeteros, disminuyendo la contaminación del suelo y el agua, y promoviendo una caficultura más sostenible y respetuosa con la biodiversidad local. Esto contribuirá a la salud del agrosistema y a la seguridad alimentaria a largo plazo.

9. Referencias Bibliográficas

- Amudha S., Bindu G., Sasi Rekha Sankar, Poongothai E., V. K., & Sarveshwaran V. (2023). Deep Learning for Plant Disease Detection and Crop Yield Prediction based on NPP-WPF Analysis in Smart Agriculture. *Preprint*.
- Araaf, R. T., Minn, A., & Ahamed, T. (2024). Coffee Leaf Rust Disease Detection and Implementation of an Edge Device for Pruning Infected Leaves via Deep Learning Algorithms. *Preprint*.
- Chatla Subbarayudu, & Mohan Kubendiran. (2024). A Comprehensive Survey on Machine Learning and Deep Learning Techniques for Crop Disease Prediction in Smart Agriculture. *Preprint*.
- Irfan, E. F., Zaka, E. R., Rehman, E. S., Sattar, B., Haider, S. A., & Hayat, M. A. (2025). An IOT-Driven Smart Agriculture Framework for Precision Farming, Resource Optimization, and Crop Health Monitoring. *Preprint*.
- Kundu, N., Rani, G., Dhaka, V., Gupta, K., Nayak, S. C., Vocaturo, E., & Zumpano, E. (2022). Disease detection, severity prediction, and crop loss estimation in MaizeCrop using deep learning. *Preprint*.
- Raveena S., & S. R. (2023). Clustering-based Hemileia Vastatrix Disease Prediction in Coffee Leaf using Deep Belief Network. *Preprint*.

-
- Tatiya, D. M., Choudhari, P., Suneel, D. A. S., Soni, D. S., Saxena, D. V., & Gore, D. R. (2025). Integration Of Smart Sensors, AI And IoT In Precision Agriculture: Advancing Crop Productivity And Sustainability. *Preprint*.
 - Thakur, D., Gera, T., Aggarwal, A., Verma, M., Kaur, M., Singh, D., & Amoon, M. (2024). SUNet: Coffee Leaf Disease Detection Using Hybrid Deep Learning Model. *Preprint*.