

Edwin Alexander Ibarra Ortiz

Contenido

Marco teórico.....	3
Justificación Metodológica del Proyecto	3
Comparación con soluciones similares.....	4
Resultados y evaluación.....	5
Impacto esperado.....	5
Propuestas de mejora	6
Reflexión ética y riesgos.....	7
Conclusión	8

Marco teórico

El avance tecnológico ha permitido desarrollar sistemas capaces de generar contenido textual especializado, esto gracias a arquitecturas como transformers modelos autorregresivos y técnicas de ajuste fino como LoRa. En particular, los modelos de lenguaje de gran escala (LLMs) han demostrado alta capacidad para sintetizar información, responder preguntas complejas y adaptarse a dominios específicos cuando se les entrena con datos curados.

En el sector agrícola la transferencia de conocimiento técnico sigue siendo un desafío. En el caso del café, aunque instituciones como Cenicafe (Centro Nacional de Investigaciones de Café) generan investigaciones y guías de alto valor, no siempre están disponibles de manera inmediata o accesible para los caficultores. En este punto es donde considero que los agentes conversacionales especializados basados en IA pueden contribuir de manera significativa en esa transferencia de conocimiento, puesto que representan una alternativa eficiente, escalable y de bajo costo.

Modelos como DeepSeek R1 (1.5B parámetros) y Llama 3.1 (8B parámetros) permiten realizar fine-tuning con técnicas como LoRA, reduciendo el costo computacional y facilitando la adaptación a dominios específicos. Estas técnicas permiten entrenar modelos de forma eficiente sin modificar todos los parámetros, lo que hace viables los proyectos de personalización incluso con recursos limitados.

Justificación Metodológica del Proyecto

El presente proyecto tuvo como meta el desarrollo de un sistema generativo destinado a proporcionar asistencia técnica e información veraz sobre el cultivo del café a los caficultores colombianos. La metodología implementada se articuló en los siguientes puntos clave:

- Selección del Dominio: Se eligió el cultivo de café reconociendo su trascendencia económica, social y cultural en Colombia, y la necesidad crítica de garantizar un acceso rápido y confiable a la asistencia técnica especializada.
- Fuente de Datos: Para asegurar la calidad, relevancia y validación científica del corpus, se empleó contenido especializado proveniente de Cenicafe, garantizando así su pertinencia en los escenarios reales del sector cafetero.
- Entrenamiento y Ajuste del Modelo: Se optó por la técnica de *fine-tuning* con LoRA. Esta elección minimizó los costos computacionales y permitió una adaptación veloz y efectiva del modelo.
 - Modelos Entrenados: El proceso incluyó el ajuste fino de cuatro modelos:

- Tres modelos DeepSeek R1 multilenguaje (1.5B), variando las épocas y los parámetros de optimización.
- Un modelo Llama 3.1 (8B) con un proceso de ajuste similar.
- Criterio de Selección: La elección de estos modelos se basó en su accesibilidad, el soporte activo de la comunidad y su eficiencia operativa, aspectos fundamentales para su despliegue en entornos con recursos limitados.
- Arquitectura y Despliegue: Se implementó una interfaz funcional y accesible públicamente mediante la herramienta Gradio. Este agente conversacional integra el modelo generativo para recibir las consultas del usuario, procesarlas y generar respuestas inmediatas.
- Impacto Social y Relevancia: La solución propuesta busca activamente reducir la brecha tecnológica y democratizar el acceso al conocimiento técnico para los pequeños y medianos caficultores, contribuyendo así a la equidad y a la mejora de la eficiencia en toda la cadena productiva del café.

Comparación con soluciones similares

Aunque la tecnología de chatbots ha llegado a la agricultura, la oferta actual de chatbots agrícolas genéricos es limitada por su falta de especialización y aplicabilidad local. Si bien existen chatbots especializados para café que cubren todo el ciclo del cultivo, su implementación en un contexto tan diverso como Colombia requiere una validación crítica y una hiper-localización que a menudo omiten.

Los puntos cruciales de adaptación necesarios son:

1. Especialización Regional: Las prácticas agronómicas y variedades varían drásticamente entre las zonas cafeteras de Colombia.
2. Temporadas: Los calendarios de siembra y floración no son uniformes y deben ajustarse al microclima.
3. Clima: Las recomendaciones (riego, sombra) deben adaptarse dinámicamente a la climatología local (temperatura, lluvia, altitud).
4. Plagas Endémicas: El manejo de enfermedades y plagas (Roya, Broca) debe ser específico para la virulencia de cada región.
5. Suelos: La recomendación de fertilización precisa requiere integrar el análisis de pH y nutrientes del suelo, lo cual los sistemas generalistas no ofrecen.

Finalmente, un gran bloqueante socioeconómico es el alto costo de suscripción o licencias de muchos sistemas especializados. Este factor crea una barrera de acceso para los pequeños y medianos caficultores, que son mayoría en Colombia, desvirtuando el sentido social y de democratización tecnológica. Toda solución debe priorizar la asequibilidad.

Resultados y evaluación

Selección y Evaluación del Modelo de Lenguaje

Se realizó una evaluación comparativa de dos modelos de lenguaje, DeepSeek R1 (1.5B) y Llama 3.1 (8B), para seleccionar el más adecuado para el ajuste fino.

Resultados de la Evaluación:

Modelo	Observaciones Clave	Evaluación de Calidad
DeepSeek R1 (1.5B)	Aumento del <i>overfitting</i> con más épocas. Respuestas inconsistentes y con mezcla de idiomas.	Baja calidad técnica según juicios expertos.
Llama 3.1 (8B)	Alta estabilidad y precisión. Demostró capacidad para sintetizar y explicar procesos agronómicos.	Alta pertinencia técnica y claridad, confirmada por evaluaciones cualitativas.

Métricas de Evaluación Aplicadas:

- Evaluación Humana: Juicios de expertos para determinar la pertinencia técnica de las respuestas.
- Coherencia Lingüística: Comparación del nivel de cohesión y fluidez del lenguaje.
- Detección de Errores: Medición de errores conceptuales y alucinaciones.

Modelo Final Seleccionado:

El modelo elegido para la implementación final fue Llama 3.1 (8B), al cual se le aplicó un proceso de *fine-tuning* utilizando la técnica LoRA, basándose en su rendimiento superior en todas las métricas.

Impacto esperado

Este proyecto es una estrategia integral con un impacto significativo y multifacético en la agricultura, centrada en el pequeño productor y la sostenibilidad.

Ejes Clave:

1. **Mejora de la Productividad:** Ofrece acceso rápido a recomendaciones técnicas basadas en datos, optimizando el uso de insumos, mejorando el manejo de plagas/enfermedades e incrementando el rendimiento y la calidad de las cosechas.
2. **Inclusión Digital y Equidad:** Funciona como motor de inclusión digital al ser una herramienta gratuita de alto valor para pequeños productores, democratizando el conocimiento técnico.
3. **Transferencia Tecnológica Efectiva:** Transforma contenido académico en conocimiento práctico, didáctico y accesible, llevando la investigación directamente al campo para una adopción rápida de innovaciones.
4. **Sostenibilidad y Resiliencia Ambiental:** Fomenta la adopción de buenas prácticas agrícolas basadas en ciencia, promoviendo la conservación de suelos, el manejo hídrico responsable y la reducción de la huella de carbono.

Potencial y Escalabilidad:

Su arquitectura permite una alta implementación y escalabilidad a través de:

- **Cooperativas Agrícolas:** Como servicio de valor añadido para asesoría técnica.
- **Gremios de Productores:** Para estandarizar la asistencia técnica regional.
- **Programas de Extensionismo:** Como base tecnológica, optimizando el trabajo de extensionistas con datos en tiempo real.

En resumen, el proyecto es un catalizador para una agricultura rural más productiva, equitativa y sostenible.

Propuestas de mejora

- No basta simplemente con tener la información publicada en cenicafé. La clave para el éxito del entrenamiento reside en la curación y preparación meticulosa del contenido. Se requiere de un equipo de personal técnico altamente experimentado con experiencia en el dominio (en este caso, el sector cafetero). La experiencia actual subraya la magnitud del reto: se realizó un ejercicio de entrenamiento con una muestra muy limitada de solo 508 registros. Si bien esto proporcionó una prueba de concepto, también dejó en evidencia numerosos frentes que deben ser atacados y mejorados para escalar la solución.
- Para este caso el mejor modelo fue el de Llama 3.1 de 8B de parámetros y el peor modelo fue de 1.5B de parámetros, vale la pena probar con un punto medio y de esta manera poder validar el modelo con menos recursos
- Validando en diversas plataformas de datasets (conjuntos de datos), se logró identificar y acceder a un recurso de gran valor: una colección de imágenes directamente relacionada con la enfermedad de la roya. Este hallazgo resulta fundamental para el proyecto, ya que

abre la puerta a la exploración e implementación de modelos de aprendizaje automático más avanzados. Específicamente, se considera de vital importancia validar la aplicabilidad de modelos multimodales. Estos modelos tienen la capacidad inherente de procesar e interpretar información proveniente de múltiples fuentes simultáneamente, lo que en este caso implicaría la habilidad de leer, analizar y extraer características relevantes de las imágenes de la roya.

Reflexión ética y riesgos

La información utilizada para el entrenamiento del modelo provino de recursos públicos disponibles en el portal oficial de Cenicafé (cenicafe.org), incluyendo boletines técnicos, guías agronómicas, fichas de recomendaciones y documentos de investigación. Estos materiales contienen lineamientos científicos validados y desarrollados por el Centro Nacional de Investigaciones de Café, entidad referente en el ámbito cafetero colombiano.

Sin embargo, aunque los contenidos se encuentran publicados de manera abierta para consulta, es fundamental considerar que su uso para fines de entrenamiento de modelos generativos puede implicar restricciones legales y éticas. Por ello, es recomendable realizar un proceso formal de validación o solicitar los permisos correspondientes a Cenicafé para garantizar que su utilización cumple con las políticas de propiedad intelectual, derechos de autor y condiciones de uso establecidas por la institución. Esto asegura transparencia, respeto por la autoría y sostenibilidad del proyecto a largo plazo, además de abrir la puerta a posibles colaboraciones futuras con la entidad.

Limitaciones del Modelo:

1. Sesgos del Corpus: El modelo podría incorporar sesgos del *corpus* de entrenamiento, lo que podría llevar a que se limite a recomendar prácticas específicas de ciertas regiones o variedades de café.
2. Dependencia Exclusiva de Cenicafé: Podría basarse únicamente en el contenido de Cenicafé, ignorando o desatendiendo importantes variaciones y prácticas a nivel local.

Estrategias de Mitigación para las Limitaciones:

- Validación Humana Continua: Incorporar la validación y supervisión continua por parte de agrónomos.

Riesgos Principales:

1. Desinformación Técnica: Existe un riesgo inherente de que el modelo genere respuestas incorrectas o técnicamente equivocadas.
 - Mitigación: Incluir advertencias claras sobre la naturaleza del consejo, realizar

validaciones periódicas y mantener el *corpus* de información actualizado constantemente.

2. Dependencia Tecnológica Excesiva: Los productores podrían desarrollar una dependencia excesiva en las recomendaciones de la Inteligencia Artificial.

Mitigación: Promover el uso de la herramienta como un complemento a la asistencia agronómica tradicional, no como un sustituto de la misma.

Conclusión

El proyecto demuestra que los modelos generativos ajustados con técnicas eficientes como LoRA pueden convertirse en herramientas de alto valor social. A pesar de los resultados insuficientes con DeepSeek R1, el desempeño del modelo Llama 3.1 evidencia que una IA generativa especializada puede proporcionar soporte técnico confiable a caficultores colombianos, democratizando el acceso al conocimiento científico y fortaleciendo la productividad del sector cafetero.