

Evaluación 02 - Desarrollo de Modelo LLM/ Agente LangChain

Jorge Iván Padilla Buriticá
Universidad EAFIT

Agosto 19, 2025

Instrucciones de la Evaluación

Esta evaluación tiene como objetivo valorar sus habilidades en el desarrollo e implementación de Modelos de Lenguaje Grandes (LLM) y agentes inteligentes utilizando herramientas de código abierto y plataformas gratuitas.

Especificaciones/Condiciones: Desarrollo y Sustentación de un Modelo LLM

Se solicita la entrega de un modelo LLM funcional desarrollado bajo las siguientes especificaciones:

- **Plataforma de Despliegue:** El modelo debe ser accesible a través de una aplicación web desarrollada con **Streamlit**.
- **Control de Versiones:** El código fuente completo del proyecto (incluyendo el modelo, la aplicación Streamlit y cualquier script de soporte) debe estar alojado en un repositorio público de **GitHub**.
- **Modelos y Herramientas:** Se deben utilizar **modelos LLM y/o herramientas de Langchain que sean gratuitos y de libre acceso**. Se recomienda explorar opciones como Hugging Face models (ej. modelos de la comunidad como 'distilbert', 'gpt2', o modelos de la serie 'Llama' si hay versiones comunitarias accesibles y eficientes para un entorno local/gratuito), o integrar APIs de modelos gratuitos si aplica. Revisar Groq.
- **Sustentación:** Se deberá realizar una presentación y demostración del modelo desarrollado con una duración máxima de **5 minutos**. Durante la sustentación, se valorará la claridad en la explicación de la arquitectura, las decisiones de diseño, la funcionalidad del modelo y los desafíos enfrentados.

Parte 1: Desarrollo de Agentes LLM (40 %)

Desarrolle **un agente LLM específico** utilizando herramientas gratuitas. Deberá elegir **una de las siguientes áreas temáticas**:

- **Biología:** Un agente capaz de responder preguntas sobre conceptos biológicos, identificar especies (a partir de descripciones textuales), o explicar procesos biológicos complejos.
- **Agricultura:** Un agente que brinde recomendaciones sobre cultivos, control de plagas, optimización de riego, o diagnóstico de enfermedades en plantas.
- **Economía:** Un agente que analice tendencias económicas, explique conceptos financieros, o simule escenarios de mercado.
- **Medio Ambiente:** Un agente que informe sobre sostenibilidad, energías renovables, impacto del cambio climático, o gestión de residuos.

Requisitos para los Agentes:

- Los agentes deben demostrar su capacidad para interactuar y resolver problemas o proporcionar información relevante dentro del dominio elegido.
- Se valorará la creatividad en la implementación y la utilidad práctica del agente.
- Se debe utilizar herramientas y modelos que sean gratuitos y de libre acceso (ej. Langchain con modelos gratuitos, o integración de APIs gratuitas si disponibles).

Parte 2: Análisis Exploratorio de Datos (EDA) e Integración con LLM (30 %)

Esta sección se enfoca en su capacidad para manejar datos, realizar un EDA (en Python) y cómo esta información puede ser utilizada para mejorar o complementar un modelo LLM.

- **Uso de Datos:** Utilice un conjunto de datos en formato **CSV** (puede ser un dataset público o generado por usted, Que tenga por lo menos 300 muestras y 6 columnas).
- **Análisis Exploratorio de Datos (EDA):** Realice un EDA significativo sobre el dataset (Mediante Python). Esto incluye:
 - Carga y visualización de los datos.
 - Identificación de valores nulos o atípicos.
 - Generación de estadísticas descriptivas.
 - Creación de visualizaciones relevantes (gráficos, histogramas, etc.) que permitan entender la estructura y las relaciones dentro de los datos.
- **Integración con Streamlit:** La aplicación Streamlit debe permitir al usuario cargar este archivo CSV y visualizar los resultados del EDA.
- **Integración con LLM:** Demuestre cómo los insights obtenidos del EDA pueden ser ingresados o utilizados por el modelo LLM desarrollado (ya sea el agente de la Parte 1 o un LLM general). Esto podría implicar:
 - Alimentar al LLM con resúmenes o conclusiones del EDA.
 - Permitir que el LLM responda preguntas basadas en el análisis de los datos.
 - Utilizar el EDA para contextualizar las respuestas del LLM.

Parte 3: Sustentación y Preguntas (30 %)

Se deberá realizar una presentación y demostración del modelo LLM y los agentes desarrollados con una duración máxima de **5 minutos**. Durante la sustentación, se valorará la claridad en la explicación de la arquitectura, las decisiones de diseño, la funcionalidad del modelo, la integración del EDA y los desafíos enfrentados.

Posibles Preguntas Técnicas y Conceptuales

Durante la sustentación, se realizarán preguntas técnicas y conceptuales relacionadas con su desarrollo. Prepárese para responder preguntas como las siguientes:

1. **Técnica:** Describa la arquitectura de su aplicación Streamlit y cómo se comunica con su modelo LLM. ¿Qué consideraciones tuvo en cuenta para optimizar el rendimiento y la experiencia del usuario?
2. **Conceptual:** Explique la diferencia entre un LLM basado en Transformers y otros modelos de lenguaje previos (ej. RNN, LSTM). ¿Cuáles son las ventajas clave de los Transformers?
3. **Técnica:** ¿Cómo manejó la integración de los modelos LLM/Langchain gratuitos en su proyecto? ¿Tuvo que realizar algún preprocesamiento o ajuste específico para el modelo?
4. **Conceptual:** ¿Qué es un **agente** en el contexto de los LLM y Langchain? Describa los componentes principales de un agente que haya desarrollado o que considere esencial.
5. **Técnica:** Detalle el proceso de despliegue de su aplicación Streamlit en un entorno gratuito (ej. Streamlit Cloud, Hugging Face Spaces). ¿Enfrentó algún desafío y cómo lo resolvió?
6. **Conceptual:** Explique el concepto de **fine-tuning** en LLM. ¿Cuándo sería apropiado aplicar fine-tuning a un modelo pre-entrenado y qué beneficios ofrece?
7. **Técnica:** ¿Cómo gestionó las dependencias y el entorno de su proyecto para asegurar la reproducibilidad del código en GitHub?

8. **Conceptual:** Hable sobre las limitaciones éticas y los posibles sesgos de los LLM. ¿Cómo considera que se pueden mitigar estos problemas en aplicaciones prácticas?
9. **Técnica:** Describa un **TOOL o Herramienta** que haya integrado o que consideraría integrar con su agente LLM para extender sus capacidades. ¿Cómo funciona la integración?
10. **Conceptual:** ¿Cuál es la importancia de la **tokenización** en el procesamiento de lenguaje natural y cómo afecta el rendimiento de un LLM?
11. **Técnica:** ¿Cómo implementó el proceso de carga de archivos CSV y la visualización del EDA en Streamlit? ¿Qué librerías utilizó?
12. **Conceptual:** ¿Cuál es la importancia del EDA antes de alimentar datos a un modelo de lenguaje o a un sistema de IA en general?
13. **Técnica:** Describa cómo los insights generados del EDA son pasados o utilizados por su modelo LLM. Proporcione un ejemplo concreto de cómo esta integración mejora la funcionalidad del LLM.

Fecha de Entrega

La fecha límite para la entrega del repositorio de GitHub (con el código del modelo LLM y los agentes) y la disponibilidad de la aplicación Streamlit es el **domingo 24 de agosto de 2025-11:59PM a través del correo jipadillab@eafit.edu.co**. Las sustentaciones se programarán para el jueves 28 de Agosto de 2025.