Esquematizar el flujo de trabajo general (pipeline)

# Pipeline LLM Deportivo

Este documento describe el flujo de trabajo general del proyecto LLM Deportivo, incluyendo sus fases, actividades y resultados esperados. Al final se adjunta un diagrama de flujo visual.

## 1️ Fase de Planeación

- Crear documento colaborativo (Google Docs)  
- Definir lista de participantes y sus roles  
  
Resultado: Organización del equipo y espacio común para documentación y seguimiento.

## 2️ Fase de Introducción

- Breve historia de los LLMs  
- Portada e introducción con definiciones clave: ¿Qué es un LLM? ¿Para qué sirve? ¿Cómo funciona?  
- Requisitos generales

## 3️ Fase de Diseño del Pipeline

### Recolección y Preparación de Datos

- Identificar fuentes confiables (APIs deportivas, FIFA, ESPN, etc.)  
- Validar calidad de datos  
- Limpiar y normalizar texto  
- Eliminar duplicados  
- Documentar fuentes y procesos de limpieza.

### Diseño y Documentación del Dataset

- Diseñar estructura clara (columnas, tipos de datos)  
- Documentar cada campo y registrar limitaciones.

### Configuración del Entorno de Desarrollo

- Elegir entorno (Colab, AWS, Azure) y justificarlo  
- Instalar dependencias necesarias  
- Validar uso de GPU  
- Crear entorno reproducible.

### Entrenamiento del Modelo

- Cargar datasets  
- Lanzar entrenamiento con monitoreo de métricas  
- Guardar checkpoints y modelos entrenados  
- Revisar longitud de prompts y calidad semántica.

### Diseño de Prompts y Contexto

- Diseñar estrategia de contexto  
- Definir instrucciones para prompts (sistema, usuario, asistente)  
- Documentar clases de tareas  
- Hacer pruebas iniciales de prompting sin fine-tuning.

### Diseño de Interfaz y UX

- Crear wireframes (Figma, Canva)  
- Diseñar flujo de usuario → modelo (input/output)  
- Desarrollar interfaz frontend y conectar al modelo  
- Documentar guía de uso para usuarios.

### Testing y Evaluación

- Crear test-set con preguntas complejas  
- Realizar pruebas UX con usuarios del equipo  
- Comparar respuestas antes/después del fine-tuning  
- Detectar errores frecuentes y recopilar feedback.

### Documentación Final y Entrega

- Documentar decisiones clave (modelo elegido, errores comunes)  
- Crear README general y README técnico  
- Registrar tiempos, parámetros, errores, mejoras propuestas.

Resultado final: Documento con el flujo completo, ordenado por fases lógicas y tareas secuenciales. Base clara para coordinar las demás actividades del proyecto LLM Deportivo.