

NLLB con LoRA : Español → Italiano

#transformers

¿Qué es un Transformer?

Un Transformer es una arquitectura que procesa secuencias de texto usando **mecanismos de atención**. En lugar de leer palabra por palabra secuencialmente, puede "mirar" todas las palabras simultáneamente y entender las relaciones entre ellas.

Componentes básicos:

- **Embeddings:** Convierten palabras en vectores numéricos
- **Atención:** Determina qué palabras son relevantes para entender cada palabra
- **Capas Feed-Forward:** Transforman las representaciones

Tipos de Transformers

Existen tres arquitecturas principales:

- **BERT (Encoder-only):** Solo "lee" y entiende. Bueno para clasificación y búsqueda, malo generando texto.
- **GPT (Decoder-only):** Solo "escribe" autocompletando. Generalista.
- **NLLB (Encoder-Decoder):** Especialista en transformación texto → texto.

Arquitectura NLLB (Encoder-Decoder)

NLLB = "No Language Left Behind"

Tiene dos módulos conectados:

1. **Encoder (Lector):** Procesa la oración completa en español y crea una representación vectorial del significado y contexto.
2. **Decoder (Traductor):** Recibe esa representación y genera la traducción en italiano token por token.

Esta separación asegura que el modelo comprenda el texto completo antes de generar, produciendo mejor gramática y orden sintáctico.

Destilación del Modelo

Usamos **NLLB-distilled-600M**:

1. El modelo original tiene **54.5B parámetros** con arquitectura Mixture-of-Experts (MoE). Demasiado grande.
2. Meta entrenó un modelo "estudiante" de 600M parámetros para imitar las salidas del modelo completo.
3. El resultado es una red densa convencional con 99% menos parámetros pero calidad comparable.

Fine-tuning con LoRA

No re-entrenamos todo el modelo. Usamos **LoRA (Low-Rank Adaptation)**:

- **Funcionamiento:** Congelamos el 99% del modelo. Solo inyectamos matrices pequeñas de parámetros entrenables en las capas de atención (q_proj, v_proj, k_proj, o_proj).
- **Ventaja:** El modelo mantiene su conocimiento lingüístico general intacto mientras se especializa en el dominio específico de tus datos.