

Informe tècnic

Per què NLLB i per què NLLB-200-600M és la millor solució per a la traducció de text offline

1. Què és NLLB?

NLLB (No Language Left Behind) és un model de traducció automàtica desenvolupat per Meta AI, dissenyat específicament per oferir traduccions multilingües d'alta qualitat. A diferència dels models de llenguatge generalistes, NLLB està entrenat exclusivament per a la traducció, amb l'objectiu de preservar fidelment el significat, el to i l'estructura del text original.

2. Avantatges principals de NLLB

a) Qualitat de traducció superior

NLLB proporciona traduccions més precises i consistents que els LLMs generalistes. No resumeix ni reinterpreta el contingut: tradueix exactament el text d'entrada.

b) Model especialitzat en traducció

Utilitza una arquitectura encoder-decoder, òptima per a tasques de traducció, fet que el fa més adequat que els models decoder-only quan l'objectiu és la fidelitat lingüística.

c) Funcionament 100% offline

Un cop descarregat el model, pot funcionar sense connexió a internet, garantint la privacitat total dels textos i la independència de serveis externs.

d) Open source i control total

NLLB és un projecte de codi obert, cosa que permet executar-lo en infraestructures pròpies, controlar els costos i complir requisits de sobirania del dada.

3. Per què NLLB-200-600M és la millor opció pràctica

La versió NLLB-200-600M (distilled) ofereix el millor equilibri entre qualitat i rendiment.

a) Equilibri qualitat / consum de recursos

Manté una qualitat de traducció molt alta amb un consum de memòria i càlcul moderat, fet que la fa ideal per a ordinadors personals i servidors locals.

b) Adequat per a macOS i entorns locals

Funciona correctament tant en CPU com amb acceleració per GPU o Apple Silicon (MPS), sense necessitat de maquinari especialitzat.

c) Cobertura lingüística suficient per a casos reals

Tot i suportar més de 200 idiomes, ofereix una qualitat excel·lent en idiomes com el català, l'espanyol i l'anglès, que són habituals en la majoria de projectes reals.

d) Superior als LLMs per a traducció offline

En comparació amb LLaMA, Mistral o Mixtral, NLLB-200-600M és més estable, més fidel al text original i més eficient.

4. Conclusió

Per a sistemes de traducció de text offline, NLLB-200-600M és la millor solució disponible actualment. Combina qualitat elevada, especialització real en traducció, funcionament offline, consum raonable de recursos i control total sobre les dades i els costos.

En resum: si l'objectiu és traduir bé, de manera fiable i sense dependre del núvol, NLLB-200-600M és l'opció òptima.