

Xinference

Xorbits Inference (Xinference) is an open-source platform to streamline the operation and integration of a wide array of AI models. With Xinference, you're empowered to run inference using any open-source LLMs, embedding models, and multimodal models either in the cloud or on your own premises, and create robust AI-driven applications.

- [Welcome to Xinference! — Xinference](#)
- GitHub: <https://github.com/xorbitsai/inference>

Revision #1

Created 16 July 2024 16:23:44 by Admin

Updated 16 July 2024 16:25:41 by Admin