

Forum:[Hack Forums](#)

Thread: [Offline Open-Source LLaMA Model - Local & Private AI, No API Fees](#)

Darth_Vadar - September 19, 2025, 3:37 pm

Hello HF,

I'm selling a pre-packaged, fully offline Open-Source LLaMA model. This isn't a subscription or API key; you get the model files to run on your own machine. This is perfect for anyone who wants a powerful AI model for their own projects without recurring costs or data going to a third-party.

Key Features:

100% Offline: No internet connection or API calls required.

Privacy-Focused: All data processing happens on your local machine.

No Recurring Costs: Pay once, own the model. No more monthly API fees.

Highly Versatile: Can be used for coding assistance, content generation, data analysis, and more.

Optimized for Performance: The model has been fine-tuned and comes with a user-friendly interface to get you started quickly.

Technical Details:

Model: LLaMA-3 based (or similar, depending on the most current stable release)

Price: OFFER ME.

This price is a one-time payment. Given the typical costs of API usage and the value of having a truly private, offline model, this is a significant long-term saving. I've seen smaller, less capable models go for a similar price on other platforms, and with the recent price hikes from big AI companies, the value of an offline solution is higher than ever.

Contact me via PM to discuss and for a demonstration. Serious inquiries only.

[contract=tid]Start Contract[/contract]

DramaKing - September 20, 2025, 5:55 am

(Sep 19, 2025 03:37 PM)Darth_Vadar Wrote: [\[CITE\]](#) Hello HF,

I'm selling a pre-packaged, fully offline Open-Source LLaMA model. This isn't a subscription or API key; you get the model files to run on your own machine. This is perfect for anyone who wants a powerful AI model for their own projects without recurring costs or data going to a third-party.

Key Features:

100% Offline: No internet connection or API calls required.

Privacy-Focused: All data processing happens on your local machine.

No Recurring Costs: Pay once, own the model. No more monthly API fees.

Highly Versatile: Can be used for coding assistance, content generation, data analysis, and more.

Optimized for Performance: The model has been fine-tuned and comes with a user-friendly interface to get you started quickly.

Technical Details:

Model: LLaMA-3 based (or similar, depending on the most current stable release)

Price: OFFER ME.

This price is a one-time payment. Given the typical costs of API usage and the value of having a truly private, offline model, this is a significant long-term saving. I've seen smaller, less capable models go for a similar price on other platforms, and with the recent price hikes from big AI companies, the value of an offline solution is higher than ever.

Contact me via PM to discuss and for a demonstration. Serious inquiries only.

[contract=tid]Start Contract[/contract]

I don't usually interfere with sales (buy requests?), but I doubt that anyone will pay for a local model.

<https://huggingface.co/meta-llama/Llama-3.1-8B-Instruct>
<https://lmstudio.ai/models/google/gemma-3-12b>

Of course, these will be quite dumb compared to frontier models, but people with recent, high-end GPUs should be able to run them.