

Fireworks / Mistral Small 3 Instruct

accounts/fireworks/models/mistral-small-24b-instruct-2501 

Serverless

LLM

Chat

[Open in Playground](#)

[Details](#) [Playground](#)

24B parameters and achieving state-of-the-art capabilities comparable to larger models

Serverless API

Mistral Small 3 Instruct is available via Fireworks' serverless API, where you pay per token. There are several ways to call the Fireworks API, including Fireworks' Python client, the REST API, or OpenAI's Python client.

See below for easy generation of calls and a description of the raw REST API for making API requests. See the [Querying text models](#) docs for details.

Try it

API Examples

Generate a model response using the chat endpoint of **mistral-small-24b-instruct-2501**. [API reference](#)

[Python](#) [Typescript](#) [Java](#) [Go](#) [Shell](#) [Chat](#) [Completion](#)



```
import requests
import json

url = "https://api.fireworks.ai/inference/v1/chat/completions"
payload = {
    "model": "accounts/fireworks/models/mistral-small-24b-instruct-2501",
    "max_tokens": 4096,
    "top_p": 1,
    "top_k": 40,
    "presence_penalty": 0,
    "frequency_penalty": 0,
    "temperature": 0.6,
    "messages": [
        {
            "role": "user",
            "content": "Hello, how are you?"
        }
    ]
}
headers = {
    "Accept": "application/json",
    "Content-Type": "application/json",
    "Authorization": "Bearer <API_KEY>"
}
requests.request("POST", url, headers=headers, data=json.dumps(payload))
```

On-demand deployments

On-demand deployments allow you to use **Mistral Small 3 Instruct** on dedicated GPUs with Fireworks' high-performance serving stack with high reliability and no rate limits.

See the [On-demand deployments](#) guide for details.

Deploy this Base model

 [Open in model playground](#)


 [Deploy this model](#)


Model Details

Created by	bchen@fireworks.ai
Created	1/30/2025
Visibility	Public
Kind	Base model
Model size	23B parameters
Function Calling	Not supported

 **Fireworks AI**







Pages	Company	Legal
Home	Blog	Trust Center
Pricing	Careers	Terms
Models		Privacy
Docs		Licenses