

Mistral AI — Modelli API per Nexus OS™

Consigliato per Nexus: `mistral-medium-3.5`

Il vero flagship di Mistral, nonostante il nome "Medium" suggerisca il contrario. Modello denso da 128B parametri, 77.6% su SWE-bench Verified, competitivo con i near-frontier di altri provider. A \$1.50/\$7.50 è il modello flagship più economico sull'output tra tutti gli 8 provider di Nexus — costa il 25% in meno di Gemini 2.5 Pro e il 50% in meno di GPT-5.6 Terra. Nella Round Table tiene testa ai competitor senza sbilanciare la fattura. Context window da 256K, reasoning configurabile, open weights (MIT modificata).

Attenzione alla nomenclatura: In Mistral la gerarchia attuale è **Medium** > **Large** > **Small** (controintuitivo). Medium 3.5 è il flagship, Large 3 è il tier economico. Non farsi ingannare dal nome.

Modelli disponibili via API — Luglio 2026

Prezzi Standard tier, per 1M token, in USD.

Flagship e modelli principali

Stringa API	Input	Output	Context	Forza principale	Note
<code>mistral-medium-3.5</code> ★	\$1.50	\$7.50	256K	Flagship: coding agentico, reasoning configurabile, vision	Consigliato — 128B denso, open weights (MIT modificata). Aprile 2026. Alias: <code>mistral-medium-latest</code>
<code>mistral-large-latest</code>	\$0.50	\$1.50	256K	Generalist affidabile, costo ultra-contenuto	Open weights (Apache 2.0). Dicembre 2025 — Alternativa economica
<code>mistral-small-4</code>	\$0.15	\$0.60	128K	Ibrido: instruct + reasoning + coding in un modello	Budget con capacità sorprendenti

Modelli specializzati

Stringa API	Input	Output	Context	Forza principale	Note
<code>codestral-latest</code>	\$0.30	\$0.90	256K	Generazione codice, fill-in-the-middle	Specializzato per IDE e coding
<code>magistral-medium-latest</code>	\$2.00	\$5.00	256K	Reasoning dedicato	Superato da Medium 3.5 per la maggior parte dei task
<code>mistral-small-3.1</code>	\$0.20	\$0.60	128K	Workhorse economico	Volume alto, task semplici

Modelli edge / ultra-economici

Stringa API	Input	Output	Context	Forza principale	Note
<code>ministral-8b-latest</code>	\$0.10	\$0.10	128K	Edge, pricing simmetrico	Ideale per classificazione e estrazione
<code>ministral-3b-latest</code>	\$0.10	\$0.10	128K	Il più veloce e economico del catalogo	244 tok/s — task molto semplici

Dati verificati sulla documentazione ufficiale Mistral — *mistral.ai* — Luglio 2026