

MiniMax-M3

KI-Modellportrait · Quellenstand 2026-10-20



MINIMAX

MiniMax-Logo (Unternehmenslogo; ein separates Modelllogo liegt nicht vor) · [Logo-Quelle / Rechteinweis](#). Marken gehören ihren jeweiligen Inhabern. Redaktionelle Identifikation, keine Partnerschaft oder Empfehlung durch den Markeninhaber.

Natives multimodales MoE-Modell mit 428 Mrd. Parametern, 23 Mrd. aktiven und 1 Mio. Token Kontext.

Quellen geprüft am 20. Oktober 2026 · Redaktion: Philipp Paulik

Entwickler: [MiniMax](#) →

Technische Einordnung

M3 wird von Beginn an gemischt-modal trainiert (Text, Bild, Video) und nutzt MiniMax Sparse Attention (MSA) für lange Kontexte. Die Modellkarte meldet 9-fache Prefill- und 15-fache Decode-Beschleunigung gegenüber M2 bei 1 Mio. Token.

Zugang & Nutzung

Gewichte auf Hugging Face und ModelScope; MiniMax Agent und MiniMax API; Inferenz über SGLang, vLLM, Transformers, KTransformers, unsloth und ATOM laut Modellkarte.

MiniMax Community License (license_name: minimax-community auf der Modellkarte); den konkreten Lizenztext im Repo vor Nutzung lesen.

Grenzen & richtig lesen

Die Tempoangaben (9x/15x) sind Herstellerangaben im Vergleich zum eigenen Vorgänger. Aktive Parameter sind kein direkter Qualitätsmaßstab. Empfohlene Inferenzparameter: temperature=1.0, top_p=0.95.

Benchmark-Belege

Für dieses Portrait ist noch kein ausreichend zugeordnetes Ergebnis erfasst. Fehlende Werte sind keine Null. Wir übernehmen keine Werte aus ähnlich benannten Modell- oder Benchmark-Versionen.

Steckbrief

Veröffentlicht

Juni 2026 (Hugging-Face-Upload 02.06.2026)

Parameter gesamt / aktiv

428 Mrd. / 23 Mrd. laut Modellkarte

Kontextfenster

1 Mio. Token

Modalitäten

Text, Bild, Video (nativ multimodal trainiert)

Architektur

MoE mit MiniMax Sparse Attention (MSA), Technical Report arXiv:2606.13392

Lizenzangabe

MiniMax Community License

Angaben laut den unten verlinkten Primärquellen.

Quellen & Stand

1. [MiniMaxAI/MiniMax-M3 \(Modellkarte und config.json\)](#)

Geprüft: 2026-10-20

2. [MiniMax Sparse Attention \(arXiv:2606.13392\)](#)

Geprüft: 2026-10-20

3. [MiniMax Group \(englischsprachiger Wikipedia-Artikel\)](#)

Geprüft: 2026-10-20

Redaktionelle Zusammenfassung mit KI-Unterstützung, verantwortet von Philipp Paulik. Kein eigener Leistungstest. Änderbare Anbieterangaben gelten für den angegebenen Quellenstand.