

腾讯混元大模型

产品简介



腾讯云

【 版权声明 】

©2013–2024 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【 商标声明 】

及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

文档目录

产品简介

产品概述

产品优势

应用场景

产品简介

产品概述

最近更新时间：2024-07-05 14:53:21

腾讯混元大模型（Tencent Hunyuan）是由腾讯研发的大语言模型，具备强大的中文创作能力，复杂语境下的逻辑推理能力，以及可靠的任务执行能力。

混元生文

混元生文目前已上线6个版本，以下为模型特性的简要介绍，可作为业务接入时版本选择的参考依据。

模型名称	能力和特征	分支版本	输入输出
hunyuan-pro	当前混元模型中效果最优版本，万亿级参数规模 MOE-32K 长文模型。在各种 benchmark 上达到绝对领先的水平，复杂指令和推理，具备复杂数学能力，支持 functioncall，在多语言翻译、金融法律医疗等领域应用重点优化。	/	最大输入 28k 最大输出 4k。
hunyuan-standard	采用更优的路由策略，同时缓解了负载均衡和专家趋同的问题。长文方面，大海捞针指标达到 99.9%。 <ul style="list-style-type: none">MOE-32K 性价比相对更高，在平衡效果、价格的同时，可对实现对长文本输入的处理。MOE-256K 在长度和效果上进一步突破，极大的扩展了可输入长度。	hunyuan-standard-32K (默认)	最大输入 30k 最大输出 2k。
		hunyuan-standard-256K	最大输入 250k 最大输出 6k。
hunyuan-lite	升级为 MOE 结构，上下文窗口为 256k，在 NLP，代码，数学，行业等多项评测集上领先众多开源模型。	/	最大输入 250k 最大输出 6k。
hunyuan-role	混元最新版角色扮演模型，混元官方精调训练推出的角色扮演模型，基于混元模型结合角色扮演场景数据集进行增训，在角色扮演场景具有更好的基础效果。	/	最大输入 4k 最大输出 4k。
hunyuan-functioncall	混元最新 MOE 架构 FunctionCall 模型，经过高质量的 FunctionCall 数据训练，上下文窗口达 32K，在多个维度的评测指标上处于领先。	/	最大输入 28k 最大输出 4k。

hunyuan-code	混元最新代码生成模型，经过 200B 高质量代码数据增训基座模型，迭代半年高质量 SFT 数据训练，上下文长窗口长度增大到 8K，五大语言代码生成自动评测指标上位居前列；五大语言 10 项考量各方面综合代码任务人工高质量评测上，性能处于第一梯队	/	最大输入 4k 最大输出 4k。
--------------	--	---	---------------------

混元多模态模型

混元多模态模型重磅发布，支持图片生文，包括图片基础识别、图片内容创作、图片多轮对话、图片知识问答、图片分析推理、图片 OCR 等能力。以下为模型特性的简要介绍，可作为业务接入时版本选择的参考依据。

模型名称	能力和特征	分支版本	输入输出
hunyuan-vision	混元最新多模态模型，支持图片+文本输入生成文本内容。 <ul style="list-style-type: none"> ● 图片基础识别：对图片中主体、元素、场景等进行识别 ● 图片内容创作：对图片进行概述、创作广告文案、朋友圈、诗词等 ● 图片多轮对话：输出单张图片进行多轮交互问答 ● 图片分析推理：对图片中逻辑关系、数学题、代码、图表进行统计分析 ● 图片知识问答：对图片包含的知识点进行问答，比如历史事件、电影海报 ● 图片 OCR：对自然生活场景、非自然场景的图片识别文字 	/	最大输入 4k 最大输出 4k。

腾讯元器平台

腾讯元器平台是一款基于混元大模型开发的一站式智能体创作与分发平台，用户无需编程经验即可轻松创建属于自己的智能体。在这个平台上，用户可以通过描述来设定智能体的角色和流程，利用丰富的插件和知识库资源让智能体更加智能化。平台还提供基于 AI 大模型的智能对话功能，以及智能体商店，汇聚了众多富有创意和趣味性的智能体玩法，为用户带来全新的智能体验。

混元生图

混元生图提供文生图 API 技术服务，基于混元文生图大模型，可结合输入的文本描述智能创作出与输入相关的图像内容，提供更高质量的图像创作、更强大的中文语义理解能力以及更广泛的应用场景。

产品优势

最近更新时间：2024-03-29 15:46:41

腾讯混元大模型基于 Transformer 神经网络架构，具有万亿参数规模。

大模型训练中，首先采用了预训练的方式，让模型对海量文本进行无监督学习，通过海量优质文本语料，让模型自动学习到自然语言的语法、语义和上下文信息（即无需给定事先标注过的训练数据，而是根据数据本身的特征和结构自动进行分类或续写等任务，目的是发现数据中的潜在规律和模式）。预训练后则进行有监督的微调，以适应特定的任务和目标。

同时，混元还采用了各种技术手段来提高模型的性能和效果，例如使用掩码策略、使用不同的优化算法、进行数据增强等。这些技术手段可帮助模型更好地处理文本数据，提高模型的泛化能力和生成效果。

腾讯混元大模型目前覆盖四大核心能力：

多轮对话

具备上下文理解和长文记忆能力，可流畅完成各专业领域的多轮对话。

内容创作

基础模型使用了中文与外文、自然语言与代码、文科与理科等知识，具备强大的中文创作能力。支持文学创作、文本概要、角色扮演能力，流畅、规范、中立、客观。

逻辑推理

准确理解用户意图、基于输入数据或信息进行推理、分析。

知识增强

有效解决事实性、时效性问题，提升内容生成效果。腾讯混元大模型采用了格式化思维链技术、“探针”算法和搜索增强技术，降低了长链条推理过程中的幻觉问题，增强了知识的实时性、真实性。

应用场景

最近更新时间：2024-03-29 15:46:41

文档场景

可提供文档创作、文本润色、文本校阅、表格公式及图表生成等能力，提高创作效率，提升创作体验。

会议场景

可提供会中问答、会议总结、会议待办项整理等能力，简化会议操作并提高会议效率。

广告场景

可提供智能化的广告素材创作，提升营销内容创作工作效率。

营销场景

构建智能导购，帮助商家提升服务质量和效率。