

澳洲5分彩官网开奖平台

EMCm7DuGMf9lBRLV

澳洲5分彩官网开奖平台当了这么久“汪峰”，阿里的Qwen3终于喜提热搜了。

当了这么多年 AI 界的汪峰，阿里 Qwen3 这回终于上了一把热搜。

距离 2.5 发布才过去 7 个月，就在今天凌晨，千问又掏出来了全新的开源全家桶，包含六款 Dense（稠密）模型和两款 MoE（混合专家）模型，能支持 119 种语言和方言。

比起 Qwen2.5 最高 72B 的参数量，千问这回的旗舰模型 Qwen3-235B-A22B 打出超级加倍，总参数量达到了庞大的 235B。

根据官方放出来的测试结果，Qwen3 在多个测试集上的表现毫不逊色国内外主流大模型，尤其是在代码和数学方面略胜一筹。

具体它的表现如何，我们也上手小测了一波旗舰模型 Qwen3-235B-A22B。

总的来说，使用体验很不错，而且在针对深度思考功能的设计上还有一些小巧思。

之前大家总嫌大模型一加深度思考就根本停不下来，想的时间太久，给的答案太细。但是不加深度思考嘛，答案的质量上又差点意思。

这回 Qwen3 把指挥棒交到用户手里，你让它想到啥程度都可以，大大提高了模型的灵活性。

不过，简单题让它简单想，难题还是得让它多琢磨琢磨。我们在测试中发现，不同的思考长度对模型的表现影响还是很明显的。

举个栗子，为了测试它的代码能力，我们想让 Qwen3 写个小游戏。

给出的提示词很简单直接，让它写一个网页上的俄罗斯方块。其他各种游戏玩法、交互、美术相关的细节，那不是人类该操心的事，让千问通过深度思考自己解决去。

而当思考长度设置在 1024 token 的时候，千问像个刚开始学代码的清澈大学生。给出的程序存在少量 bug，根本玩不起来。

但预算拉满之后，它成了熟练的老码农，只花几分钟就能搓出来一个完全体俄罗斯方块。

接下来，我们让中文互联网上难度最高深莫测的逻辑测试集开始表演：

“平时烧水很麻烦，为什么不一次性烧好多水然后冻起来，等需要的时候再拿出来呢？”

在关闭深度思考的时候，模型还会一本正经地胡说八道：

《节省时间》、《节能》、《确实方便》，说得这么有理有据，我信了。

而一旦启动深度思考，模型一眼看出来这就是个奇葩问题，直接对逻辑提出异议。

前段时间，OpenAI 在 o3 的官方文档中就表示，它们发现模型的推理时间越长，效果越好。

而 Qwen3 的这些个例子算是证明了，通过更长时间的深度思考，大模型确实智商猛涨。

另外，既然代码和逻辑都难不住它，那就再试试千问在多模态上的表现咋样。

前一阵子 GPT-o3 的图片推理都让大家伙儿脊背一凉，这次大升级的 Qwen3 也会成为开盒神器吗？

会的兄弟，会的。

有的差友可能还记得，前不久我们做了一期 o3 开盒，它靠着民宿的招牌定位到了梦想小镇。

这回 Qwen3 更离谱，下面这张照片里没有一个字，你知道它是用什么验证猜测的吗？

没错，是照片左侧的一个爱心雕塑。怕大家看不出来，我特意在上面用红框圈了一下，没注意的差友可以再仔细找找。

这回不能说人家靠照片内置信息作弊了，千问开盒和马斯克的智驾一样，纯视觉。

除了以上这些传统艺能，Qwen3 还追上了 MCP 的热潮。虽然目前还在测试中没有开放，但官方秀出了两个案例。

给它一个 Github 库，千问可以自己浏览查询网页上的信息，总结每个项目的 star 数，再画出柱状图。

让它分类归纳乱乱的桌面文件也是小菜一碟。

先帮你把文件夹创建好，然后一秒合并同类项，纵享丝滑。

我只想说：这些功能什么时候上线？自动收集数据作图是真实存在的吗，摸鱼星人狠狠心动了！

测试看完了，有的差友可能对 Qwen3 的技术细节还有点疑惑：它到底跟之前的大模型有啥区别？

简单来说，之前的大模型，推理和快速回答都是分开的。比如 DeepSeek-R1 和 GPT-o3 属于推理模型，而 DeepSeek-V3 和 GPT-4o 负责快速响应。

现在的 Qwen3-235B-A22B，则是一个“混合推理模型”，相当于 R1+V3，o3+4o。

但模型加功能可不是做个加法这么简单。Qwen3 具体是怎么训练出来这个二合一模型的呢？

官方在文档里展示的后训练四步走，解答了这个问题。

谜底就在 post-training 第三阶段，Qwen3 把长思维链的数据和普通的常用指令放在一起对模型进行了微调。

这样就可以把快速回答模式整合到深度思考模型中，确保推理和快速响应能力的无缝结合。

目前，这种混合大模型，国外有个闭源 Claude 3.7 Sonnet，而国内只千问一家，团队还把它慷慨开源了！

说到这里，有的小伙伴肯定已经在摩拳擦掌准备冲了。

但这 235B 的参数量是不是看起来有点。。。

别慌，MoE 模型有一个大大的好处就是，回答问题不用全员上阵，大部分员工都在休息，所以每次激活的参数量并不大，只有 22B，而真正吃性能的，也只有这 22B。

也就是说，速度更快，成本更低了。官方表示，部署 Qwen3-235B-A22B 只需要 DeepSeek-R1 35% 的成本。

而 Qwen3 刚一上线也是备受关注，迅速攀升 Hugging Face 热搜榜。

这回千问团队同样提供了不同规模的蒸馏模型，一共 8 款任君挑选，最小的 0.6B 模型在移动端都能跑，总有一个符合你的需求。

我们也把 0.6B 的版本的 Qwen3，迅速部署到了手机上，试用了一下，效果还挺乐的：

不过这已经是最小的模型了嘛，要求不能太高，至少好玩。

总的来说，这一次的 Qwen3 更新，又给大模型的开源圈带来了一大波狠货。

Qwen 在大模型开源圈儿的地位，也进一步得到了巩固，按照阿里云官方说法，在开源圈发育了这么久，目前千问的衍生模型已经超 10 万个，全球下载量超 3 了亿次，甚至把之前的开源第一 Llama 系列都甩在了后头。

甚至在某种程度上，AI 圈处处都有千问的影子。

比如，为啥叫千问 AI 圈汪峰呢？因为它每次出新品的时候，总被更狠的活儿压下去。

Qwen2.5-Max 撞了 DeepSeek-R1，3 月 QwQ-32B 又撞 Manus。

但其实，DeepSeek-R1 论文中的蒸馏模型案例，是通过千问和 Llama 整的；Manus 的创始人也公开表示，他们的产品也用了是在千问的基础上微调开发的。

所以，虽然这个热搜迟到了，但通义千问在国产大模型的发展历程中，其实一直没咋缺席。

最后，求求 DeepSeek 再加个速吧，R2 已经等不及辣！

撰文：莫莫莫甜甜

编辑：江江 & 面线

美编：焕妍

图片、资料来源：

莫莫实测、不咕实测

<http://qwenlm.github.io/blog/qwen3/> <https://github.com/QwenLM/Qwen3>

<https://www.usenix.org/system/files/atc23-zhai.pdf>

<https://huggingface.co/> <https://ollama.com/library/deepseek-r1:8b>

<https://arxiv.org/pdf/2501.12948>

<https://baijiahao.baidu.com/s?id=1826262065502821151&wfr=spider&for=pc>
<https://openai.com/index/introducing-o3-and-o4-mini/>

澳洲10冠军定码

澳洲幸运10官方开奖结果

168澳洲幸运5官方开奖记录查询

澳洲计划免费一期计划

全天计划系统

澳洲10二期计划

澳洲幸运5开奖结果 开奖结果

澳洲5分开奖官网开奖

澳洲10开奖记录岸

澳洲幸运10赢怎么看走势

重庆水果彩票幸运农场

一分钟极速赛车开奖结果

澳洲10会控制吗

计划之家app官网下载

极速赛车168开奖记录一分钟

澳洲10开奖记录科

澳洲幸运5开奖直播软件

幸运飞行艇官网app

澳洲10二期计划