

2025澳洲幸运5官方开奖

EMCm7DuGMf9lBRLV

2025澳洲幸运5官方开奖Qwen 3发布，阿里又点燃了AI开源的篝火

文 | 邓咏仪

编辑 | 苏建勋

4月28日，AI圈子内的从业者们都在等待一件事：Qwen 3。

从中午开始，Qwen 3即将发布的小道消息，已经满天飞舞。Qwen团队负责人林俊暘也在X上暗示：“看看我们今晚能否完成Qwen 3的工作”。

来源：X (Twitter)

《智能涌现》所在的多个业内讨论群，充斥着不知真假的Qwen 3模型上传截图。AI从业者疯狂刷新GitHub、HuggingFace中的Qwen主页，用AI生成Qwen 3上线海报、现场的模拟图，刷屏各种表情包，狂欢直至深夜还未停歇。

Qwen 3最终在凌晨5点上线。新一代的Qwen 3参数量仅为DeepSeek-R1的1/3，首先是成本大幅下降，性能全面超越R1、OpenAI-o1等全球顶尖模型。

更重要的是，Qwen 3搭载了Claude 3.7等顶尖模型的混合推理机制，将“快思考”与“慢思考”集成进同一个模型，大大减少了算力消耗。

Qwen 3的开源一共涉及8款不同架构和尺寸的模型，从0.6B开始，大到235B，适用于更多类型的移动端设备。除了模型之外，Qwen还顺带推出了Agents的原生框架，支持MCP协议，有着一股“让所有人都要用上Agents”的劲头。

DeepSeek爆火后的1月份，大年初一前夜，阿里急速上线了新模型Qwen2.5-VL和Qwen2.5-Max，迅速秀了一把肌肉的同时，也赋予了阿里集团更浓郁的“AI味儿”。受此情绪的烘托，阿里股价在春节前后阶段大涨超30%。

但比起这次的旗舰级模型Qwen 3，上述模型都只是前奏。

对Qwen 3的万众期待，来自阿里在AI开源社区的盛誉——如今Qwen已经是全球领先的开源模型系列，根据最新数据，阿里通义已开源200余个模型，全球下载量超3亿次，千问衍生模型数超10万个，已超越此前的开源霸主Llama。

如果说DeepSeek是一支精锐的小分队，快速在技术上冲锋；那么Qwen就是一个军团，对大模型布局更早，也更积极做生态，展现了更广的覆盖度和社区活力。

某种程度上，Qwen也是大模型落地的产业风向标。

一个典型的例子是，在DeepSeek R1发布后，很多企业和个人想私有化部署“满血版”DeepSeek (671B)，单单硬件成本就要上百万元，落地成本很高。

阿里Qwen家族提供了更多的模型尺寸和类别，能帮助产业界更快验证落地价值。用大白话来讲，开发者不用自己裁剪模型，而是拿来即用，继而快速落地。Qwen 1.5B及以下的模型，可控性强，的确是如今AI应用领域最受欢迎的模型之一。

DeepSeek R1成为开源历史上的锚点，也深深影响了大模型竞争的走向。不同于此前模型厂商沉迷于刷Benchmark、刷题，中国大厂们迎来必须要证明自己真正技术实力的周期。

Qwen 3的发布，就是这样一个时点。

满血版成本为1/3个DeepSeek R1，性能更强大

2024年9月，阿里云在云栖大会上发布了上一代模型Qwen 2.5。Qwen2.5 一次性开源了从0.5B至72B共6个尺寸的全系列模型，覆盖从端侧到云端的全场景需求，在代码等多个类别中，都能达到SOTA。

所有模型均允许商业使用与二次开发，这也被开发者称为“真正开放的AI”。

市场风传新一代的Qwen 3会推出基于MoE架构，开源更多尺寸，成本能做到比DeepSeek更低——这些猜测都被一一证实。

Qwen 3 一共开源了8个尺寸的模型，分别为：

每款模型均达到了同尺寸开源模型中的SOTA（最佳性能）。

Qwen 3延续了阿里开源的慷慨风格，依旧采用宽松的Apache2.0协议开源，首次支持119多种语言，全球开发者、研究机构和企业，都免费下载模型并商用。

Qwen 3的最大看点，一是成本大幅下降，同时性能还有大幅提高。

在Qwen 3的训练中，阿里投入了令人惊讶的成本。Qwen 3基于36万亿token进行预训练，这是上一代模型Qwen 2.5的两倍，在全球同等体量的顶尖模型中也能排到前列。

据Qwen团队公开的数据，仅需4张H20即可部署Qwen 3满血版，显存占用仅为性能相近模型的三分之一。

成本降低，但性能反而更高了。

Qwen的推理能力有非常显著的提升。在数学、代码生成和常识逻辑推理方面，Qwen 3均超越先前的推理模型 QwQ（思考模式）和 Qwen2.5 模型（非思考模式）。

而在代码、数学、通用能力等基准测试中，Qwen 3也能与目前的顶尖模型o3-mini、Grok-3 和 Gemini-2.5-Pro等顶尖模型。

△Qwen 3性能图 来源：Qwen 3

另外一个核心亮点，是Qwen模型对智能体（Agents）的全面适配。

如果说，OpenAI的o1模型一脚踢开了推理模型的大门，DeepSeek R1的发布，让所有用户见识到了推理模型的魔力：模型有像人类一样的“思考链”，有思考顺序，不断验证是否正确，推演出认为合适的答案。

但只有深度思考模式的话，弊端也很明显。哪怕询问简单的天气、今天穿什么，DeepSeek会来回地纠结，自我设问，不断验证，经历少则数十秒的过度思考——如果DeepSeek没有将模型的思考链展示给用户，几乎没有用户能够忍受延迟成这样的对话体验。

阿里CEO吴泳铭曾在2024年9月的云栖大会表示：“AI最大想象力不在手机屏幕，而是接管数字世界，改变物理世界”。

智能体是通往这种愿景的重要路径。所以，Qwen3做成混合推理模型是相当关键的：在单一模型内，能无缝切换思考模式（用于复杂逻辑推理、数学和编码）和非思考模式（用于高效的通用对话，比如询问天气、历史知识等简单信息搜索）。

推理与非推理任务的融合能力，实际上是让模型：

Qwen 3的API可按需设置“思考预算”（即预期最大深度思考的tokens数量），进行不同程度的思考，确保在各种场景下都能达到最佳性能。

在以前的机制中，用户需要手动开关“深度思考”模式，一次对话中可能只能专注一种模式；但Qwen 3的新机制将这个选择交给模型——模型能自动识别任务场景、选择思考模式，减少了用户对模型模式的干预成本，也会带来更丝滑的产品体验。

混合推理是目前比较难的技术方向，需要极其精细、创新的设计及训练，难度远超训练纯推理模型。模型要学习两种不同的输出分布，要做到两个模式融合，且基本不影响任何一种模式下的效果。

热门模型中，现在只有Qwen 3、Claude3.7以及Gemini 2.5 Flash可以做到较好的混合推理。

混合推理会整体提高模型使用的性价比，既提升了智能水平，又整体降低了算力资源消耗。比如，Gemini-2.5-Flash，推理和非推理模式下的价格相差约6倍。

而为了让所有人都能马上开发Agents，Qwen团队几乎是提供了保姆式的工具箱：

如果用装修来举例，这就像Qwen团队把房子建好、完成硬装，还给你提供了部分软装，开发者可以直接用上很多服务。这将大大降低编码复杂性，开发门槛进一步下降，比如很多手机及电脑Agent操作等任务，就可以高效实现。

开源模型进入新一轮竞争周期

在DeepSeek R1获得爆炸性声量，成为全球开源模型的标杆之后，模型发布不再是单纯的产品更新，而是代表公司战略的关键走向。

Qwen 3的发布正值DeepSeek R1后，开源社区的新一轮竞争已经开始：2025年4月，Meta旗下的Llama 4在4月初正式发布，但因为效果不佳而遭诸多批评；而此前屡屡受挫的AI巨头Google，也借着Gemini 2.5 pro，在开源领域扳回一城。

通用大模型层的能力仍在快速变迁中，很难有厂商能够一直保持领先。在这个时点上，大模型团队如何确定自己的发展主线，不只是个技术问题，更是对不同产品路线和商业判断的策略问题。

在Qwen 3的发布上，可以看出一种更务实的开源策略。

比如，Qwen3本次提出的模型尺寸，就比Qwen 2.5时的尺寸划分就更细致。在资源受限的设备（如移动端、边缘计算设备）上实现高效运行，Qwen 3能同时保证一定的性能，满足轻量级推理、对话等需求。

阿里仔细解释了各个模型的适用场景：

而在旗舰模型上，Qwen 3的模型规模和架构，也是一个更精炼，更容易落地应用的设计。

以Qwen旗舰版模型235B（235亿参数）和DeepSeek R1满血版做直接对比：

从部署成本看，Qwen 3是满血版R1的25%-35%，模型部署成本大降六到七成。

DeepSeek R1之后，如果说大模型领域达成了什么共识，那便是——重新将资源、人力投入到模型层的技术突破，让模型能力突破应用能力的上限。

如今大模型领域，重新将目光转向模型能力的突破。

从Qwen发布主题的变化，也能看出如今技术主线的变迁：Qwen 2.5发布时，主题还是《扩展大型语言模型的边界》，而到了Qwen 3，则是《思深，行速》，专注提高模型能力的应用性能，拉低落地门槛，而非单纯扩大参数规模。

现在，通义千问Qwen在全球下载量超过3亿，在HuggingFace社区2024年全球模型下载量中，Qwen占比已经超30%。阿里云的模型开源策略，已经走出另一条更清晰的道路：真正成为应用的土壤。

欢迎交流

欢迎交流

168澳洲幸运十开奖记录

快三数据分析器软件

澳洲10怎么玩才能赢

168飞艇计划全天预测永久免费

澳洲10开奖和历史

澳洲幸运10开奖结果官方直播

澳洲十分计划

澳洲10开奖结果官网

澳洲人工费用一般多少?

168幸运飞艇官方开奖记录

1234公式准确率100

pc28一天赚2000技巧

澳洲幸运10历史开奖号码查询

澳门118期开奖

腾讯分分彩在线玩

爱奇艺怎么购买单场球赛

澳洲幸运10技巧图片

澳洲幸运5开奖结果号码查询

极速赛车168开奖官网开奖视频