

月之暗面 Moonshot AI

从万亿基座到智能涌现

深入探索 KIMI K2 & K2 Thinking 的技术原理，了解万亿参数混合专家模型如何实现智能体范式的突破

Z

1.04T

总参数量

32B

激活参数

384

MoE 专家数

256k

上下文长度

课程模块

系统化学习 KIMI K2 从架构设计到智能体训练的完整技术体系



KIMI K2 简介

了解万亿参数级混合专家大语言模型核心理念



架构设计

深入了解 1T MoE 架构与工程权衡



预训练技术

攻克稳定性与效率的创新方案



K2 Thinking

智能体训练范式的突破



推理与部署

探索能力边界与实际应用



Demo体验

实时体验 K2 Thinking 的智能对话能力

核心创新

K2 在多个技术维度实现突破性创新

Z

MuonClip 优化器

通过 QK-Clip 机制实现 15.5 万亿 Tokens 训练零损失尖峰，保证万亿模型训练的稳定性

数据改写流水线

充分利用高质量数据，通过多视角改写提升词元效率，避免讨论会

智能体数据合成

大规模模拟真实工具使用场景，训练 3000+ 真实工具和 20000+ 合成工具

原生 INT4 量化

量化感知训练实现 2 倍推理加速，显著降低显存占用且性能几乎无损

超长调用链

突破行业瓶颈，稳定完成 200-300 次连续工具调用，保持目标一致性

测试时扩展

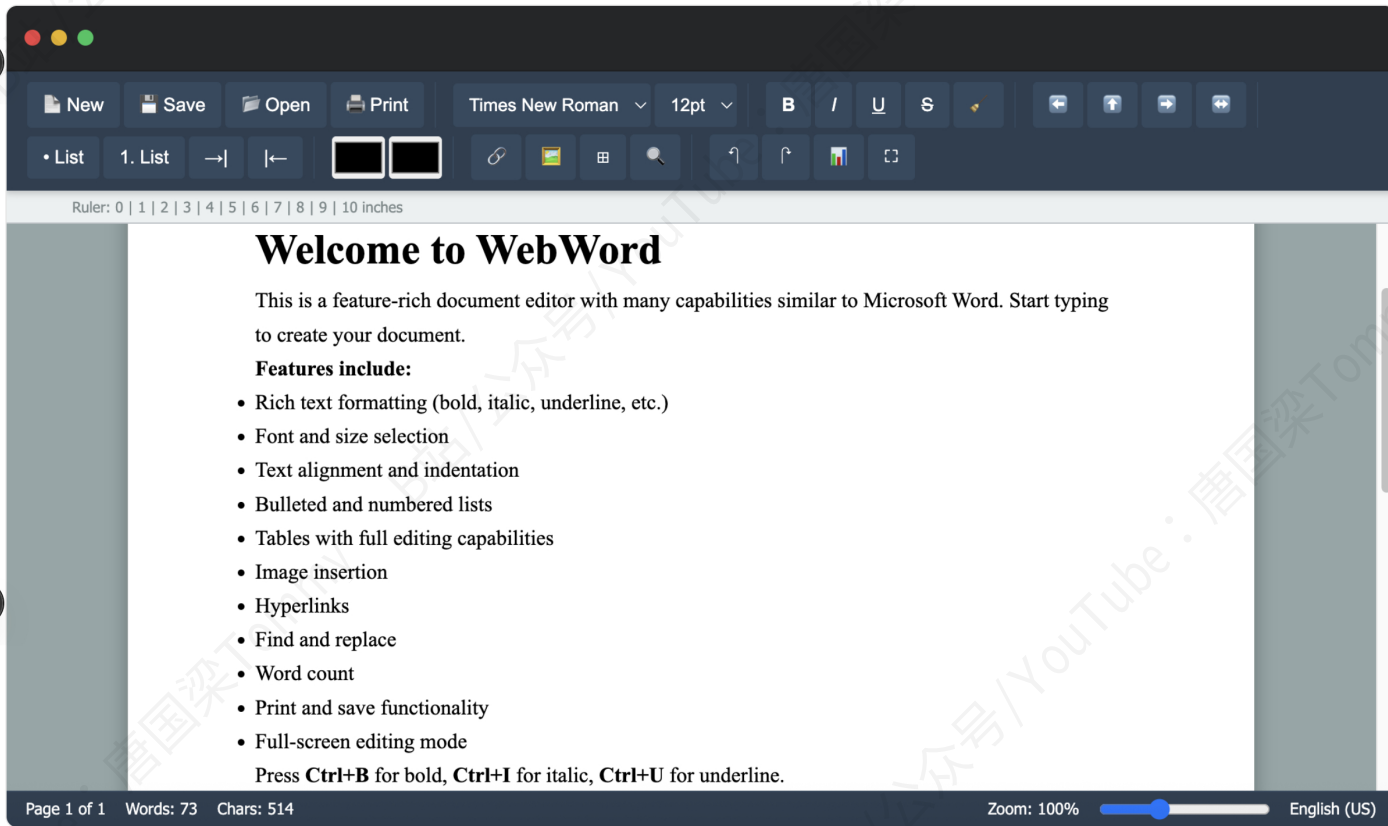
在推理时动态扩展模型能力，通过更多思考时间提升复杂任务表现

应用案例

探索 KIMI K2 Thinking 在不同领域的实际应用

案例 1: Component-heavy Website

构建复杂的组件驱动网站，展示 K2 在前端开发和组件架构设计中的能力

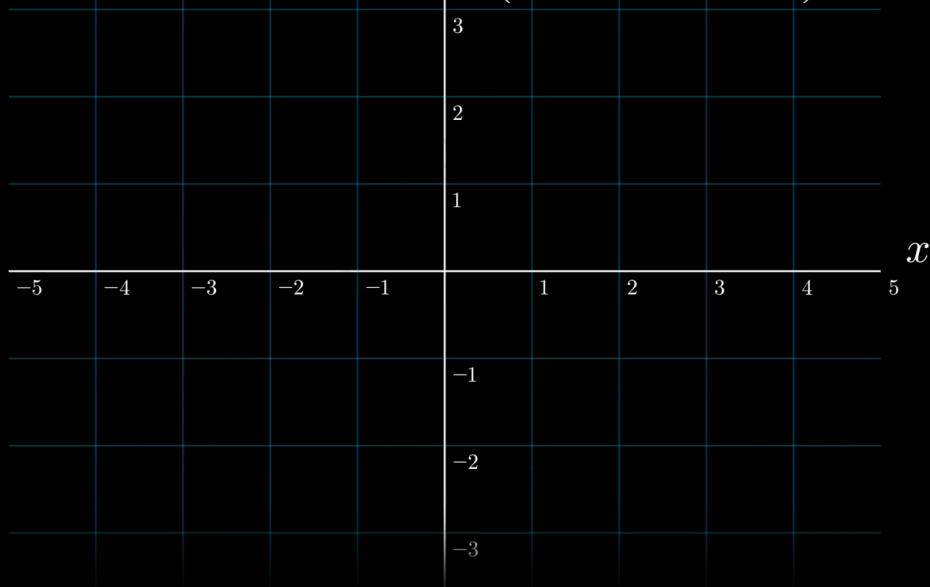


案例 2: Math Explainer

创建交互式数学解释器，结合可视化和步骤分解，帮助理解复杂数学概念

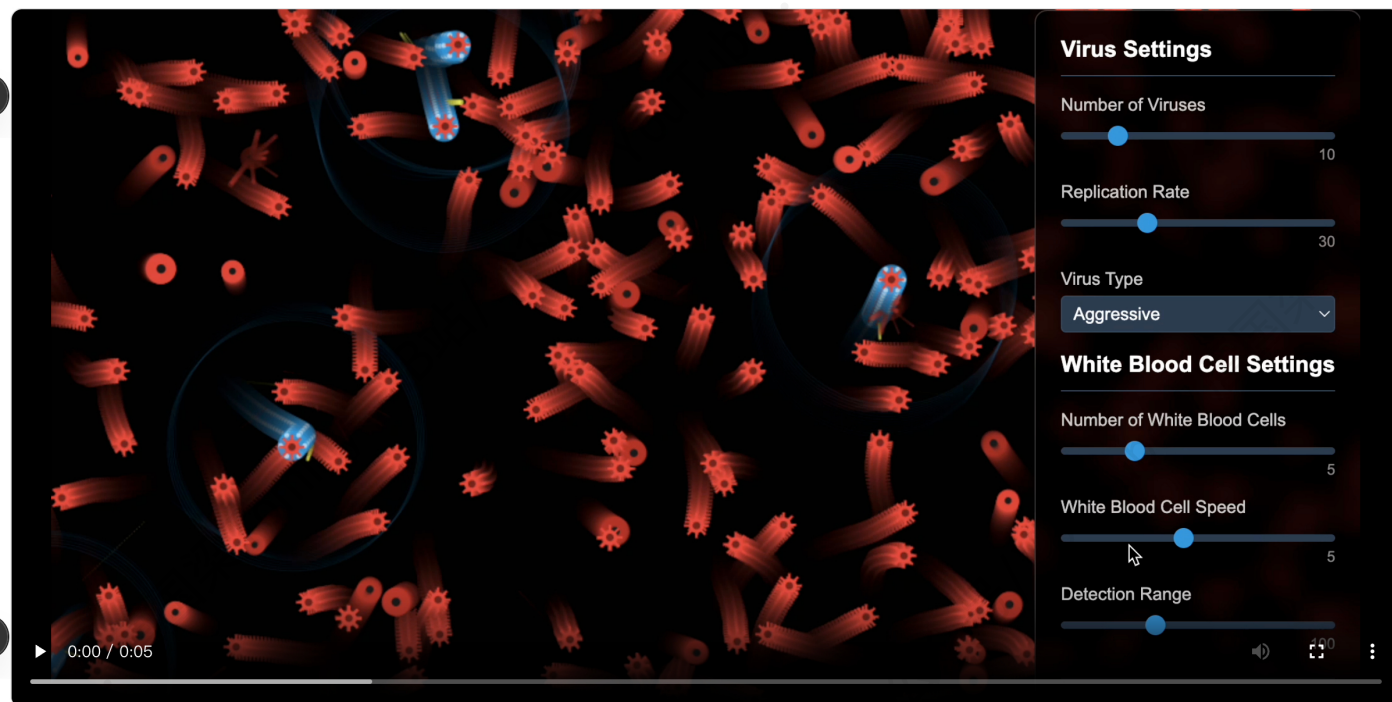


Gradient Descent in 2D (Contours + Path)



案例 3: Simulation of virus attacking cells in bloodstream

模拟病毒攻击血液细胞的生物过程，展示 K2 在科学可视化和仿真领域的应用



开始探索 KIMI K2

系统学习从万亿参数基座模型到智能体范式突破的完整技术路径

立即开始 →

2

© 2025 唐国梁Tommy