BIS - MT · 19/06/2026

Bản Tin AI Hằng Ngày

Cập nhật công nghệ AI mới nhất

"A smooth sea never made a skilled sailor."

↪ Biển lặng không tạo ra người thủy thủ giỏi.

— Franklin D. Roosevelt

Kỹ năng và bản lĩnh chỉ được rèn giũa qua thử thách — đừng tránh né khó khăn mà hãy xem chúng như cơ hội để phát triển.

TIN TỨC NỔI BẬT

1

Các nhà nghiên cứu bảo mật cảnh báo về lỗ hổng trong code do AI tạo ra

Security Researchers Sound the Alarm on Vulnerabilities in AI-Generated Code

Infosecurity Magazine Đọc bài viết →

Các nhà nghiên cứu bảo mật đã bày tỏ mối quan ngại về các điểm yếu bảo mật trong mã được tạo ra bởi AI, đang ngày càng được sử dụng trong nhiều ngành công nghiệp. Các nhà nghiên cứu cảnh báo rằng mã được tạo ra bởi AI có thể chứa các điểm yếu bảo mật ẩn, khiến nó trở thành một rủi ro tiềm ẩn đối với hệ thống và dữ liệu. Vấn đề này xuất phát từ thực tế rằng các model AI thường được đào tạo trên mã hiện có, có thể chứa các điểm yếu. Những điểm yếu này sau đó có thể được sao chép trong mã được tạo ra bởi AI, khiến nó dễ bị tấn công. Hơn nữa, mã được tạo ra bởi AI cũng có thể khó kiểm tra và thử nghiệm, khiến việc xác định và sửa lỗi bảo mật tiềm ẩn trở nên thách thức. Các nhà nghiên cứu nhấn mạnh sự cần thiết cho các developer phải nhận thức được những điểm yếu này và thực hiện các biện pháp phòng ngừa cần thiết khi sử dụng mã được tạo ra bởi AI. Họ đề xuất rằng các developer nên xem xét và thử nghiệm kỹ lưỡng mã được tạo ra bởi AI trước khi triển khai nó trong môi trường sản xuất. Các nhà nghiên cứu cũng khuyến nghị rằng các developer nên sử dụng nhiều model AI và xem xét mã được tạo ra một cách thủ công để đảm bảo tính bảo mật của nó.

2 Lớp lệnh AI kho hàng đa agent giúp tối ưu vận hành và nâng cao thông minh chuỗi cung ứng

Multi-Agent Warehouse AI Command Layer Enables Operational Excellence and Supply Chain Intelligence

NVIDIA Developer Đọc bài viết →

NVIDIA đã giới thiệu Lớp lệnh AI Nhà kho Đa tác nhân, được thiết kế để nâng cao hiệu quả hoạt động và trí tuệ chuỗi cung ứng. Hệ thống này được hỗ trợ bởi AI, cho phép giao tiếp và phối hợp liền mạch giữa các tác nhân nhà kho khác nhau, chẳng hạn như robot, máy bay không người lái và công nhân người. Bằng cách tận dụng AI, hệ thống có thể tối ưu hóa hoạt động nhà kho, dự đoán và ngăn chặn tình trạng tắc nghẽn, và cải thiện quản lý hàng tồn kho. Lớp lệnh AI Nhà kho Đa tác nhân sử dụng nền tảng Omniverse của NVIDIA, cung cấp một môi trường có thể mở rộng và bảo mật để phát triển và triển khai các ứng dụng được hỗ trợ bởi AI. Nền tảng này cho phép tạo ra các mô phỏng phức tạp, cho phép các nhà điều hành nhà kho kiểm tra và tinh chỉnh các chiến lược AI của họ trước khi triển khai. Khả năng AI của hệ thống cũng cung cấp thông tin chi tiết theo thời gian thực về hoạt động nhà kho, cho phép đưa ra quyết định dựa trên dữ liệu và cải thiện quản lý chuỗi cung ứng. Bằng cách tối ưu hóa hoạt động nhà kho và nâng cao trí tuệ chuỗi cung ứng, Lớp lệnh AI Nhà kho Đa tác nhân nhằm mục đích cải thiện hiệu quả tổng thể và giảm chi phí cho các công ty hậu cần và thương mại điện tử.

3 Giới thiệu Open Agent Specification (Agent Spec): Một chuẩn thống nhất cho AI Agent

Introducing the Open Agent Specification (Agent Spec): A Unified Representation for AI Agents

Oracle Blogs Đọc bài viết →

Oracle đã giới thiệu Open Agent Specification (Agent Spec), một biểu diễn thống nhất cho các tác nhân AI. Agent Spec nhằm cung cấp một khuôn khổ tiêu chuẩn hóa để mô tả và tương tác với các tác nhân AI trên nhiều nền tảng và ứng dụng khác nhau. Thông số kỹ thuật này được thiết kế để mã nguồn mở và do cộng đồng dẫn dắt, cho phép các nhà phát triển đóng góp và định hình sự phát triển của nó. Agent Spec được xây dựng trên các tiêu chuẩn và công nghệ hiện có, bao gồm JSON-LD và JSON Schema. Nó cung cấp một biểu diễn có cấu trúc của các tác nhân AI, cho phép giao tiếp và tích hợp liền mạch giữa các hệ thống khác nhau. Biểu diễn thống nhất này dự kiến sẽ thúc đẩy sự

phát triển của các ứng dụng và dịch vụ AI phức tạp hơn. Agent Spec không giới hạn ở một loại tác nhân AI cụ thể, chẳng hạn như chatbot hoặc trợ lý ảo, mà cung cấp một khuôn khổ chung để biểu diễn bất kỳ loại tác nhân AI nào. Bằng cách cung cấp một cách tiêu chuẩn hóa để mô tả và tương tác với các tác nhân AI, Agent Spec có tiềm năng đẩy nhanh việc áp dụng và phát triển các công nghệ AI trên nhiều ngành công nghiệp khác nhau.

4

Tại sao LLM nên ngừng "nghĩ lớn tiếng" (và điều gì đến sau chain-of-thought)

Why LLMs should stop thinking out loud (and what comes after chain-of-thought)

BD Tech Talks Đọc bài viết →

Ngành công nghệ đang đối mặt với một thách thức đáng kể khi chi tiêu token AI tăng vọt mất kiểm soát, với các công ty lớn như Uber, Meta và Amazon đang□□ đánh giá lại các quy trình và ngân sách AI của họ. Một yếu tố đóng góp chính cho vấn đề này là việc sử dụng kỹ thuật Chain-of-Thought (CoT) prompting, nơi các mô hình ngôn ngữ lớn (LLMs) được hướng dẫn để "nghĩ từng bước" trước khi đưa ra câu trả lời cuối cùng. Mặc dù CoT ban đầu là một giải pháp thực tế tận dụng các giao diện tạo văn bản hiện có, nó đã trở thành một tiêu chuẩn ngành mặc dù có những hạn chế. Nghiên cứu đã chỉ ra rằng CoT không phải là một cơ chế suy luận thực sự, mà chỉ là sự bắt chước, và rằng các token trung gian được tạo ra bởi LLMs thường hoạt động như một ràng buộc cấu trúc chứ không phải là sự phản ánh thực sự của quá trình tính toán nội bộ của mô hình. Điều này đã dẫn đến một tâm lý cargo cult, nơi việc tạo ra các token văn bản trung gian tốn kém được coi là tương đương với quá trình xử lý nhận thức thực sự. Để mở rộng AI một cách bền vững, ngành công nghệ cần phải vượt qua CoT và khám phá các cơ chế suy luận thay thế.

5

Tuần lễ Giáo dục 2026 của IEEE nhấn mạnh học tập suốt đời

IEEE's 2026 Education Week Events Emphasized Lifelong Learning

IEEE Spectrum Đọc bài viết →

Cộng đồng IEEE gần đây đã kết thúc Tuần Giáo dục thứ năm hàng năm, một sự kiện kéo dài một tuần nhấn mạnh việc học tập suốt đời. Từ ngày 11 đến 19 tháng 4, tổ chức này đã cung cấp một loạt các sự kiện trực tiếp và ảo, tài nguyên trực tuyến và khuyến mãi cho các

chuyên gia và sinh viên. Sự kiện này bắt đầu với một bài phát biểu quan trọng của Chủ tịch IEEE Mary Ellen Randall, nhấn mạnh tầm quan trọng của tài nguyên giáo dục trong các lĩnh vực STEM. Tuần lễ kỷ niệm này đã có sự tham gia của hơn 120 đối tác IEEE, sản xuất 114 sự kiện, 23 tài nguyên và 11 ưu đãi đặc biệt. Sự kiện này nhằm cung cấp các công cụ dễ tiếp cận cho các thành viên để phát triển chuyên môn và hướng dẫn các kỹ sư mới. Các chỉ số tham gia cho thấy sự quan tâm địa lý rộng rãi, với sự tham gia chính đến từ Ấn Độ, Nigeria và Hoa Kỳ. Trang web Tuần Giáo dục IEEE đã ghi nhận hơn 4.770 lượt truy cập, và gần 240 huy hiệu kỹ thuật số đã được cấp cho những người hoàn thành các bài kiểm tra giáo dục. Sự thành công của sự kiện này đã mở đường cho việc lên kế hoạch cho Tuần Giáo dục IEEE 2027, dự kiến từ ngày 3 đến 11 tháng 4.

6

Từ các dự án open source đình đám đến OpenAI

From open source hits to OpenAI

Changelog Đọc bài viết →

Tuần này, Max Stoiber, một developer tại OpenAI, tham gia một podcast để thảo luận về các dự án mã nguồn mở khác nhau và tác động của chúng đối với ngành công nghiệp công nghệ. Ông đề cập đến một số dự án đáng chú ý, bao gồm react-boilerplate và styled-components, cũng như việc GitHub mua lại Spectrum, điều này đã giúp định hình GitHub Discussions. Stoiber cũng nói về sự phát triển của Stellate, một bộ nhớ đệm GraphQL đã được Shopify và The Guild mua lại. Ngoài ra, ông thảo luận về sự phát triển của các ứng dụng ChatGPT, mà ông tin rằng cung cấp một bề mặt mới cho phần mềm. Tập này cũng có một số tài trợ, bao gồm Coder.com, WorkOS, Notion, Fly.io và Changelog++.

7

Hình dạng, đối xứng và cấu trúc: Vai trò thay đổi của toán học trong nghiên cứu Machine Learning

Shape, Symmetries, and Structure: The Changing Role of Mathematics in Machine Learning Research

The Gradient Đọc bài viết →

Trong những năm gần đây, nghiên cứu học máy đã chứng kiến sự chuyển dịch đáng kể từ các phương pháp dựa trên nguyên tắc toán học sang các phương pháp thực nghiệm và kỹ thuật hơn. Mặc dù vậy, toán học vẫn là một thành phần quan trọng của nghiên cứu học máy,

mặc dù vai trò của nó đang thay đổi. Trong khi các đảm bảo lý thuyết truyền thống về hiệu suất của model không còn là trọng tâm chính, toán học đang được áp dụng theo những cách mới, chẳng hạn như cung cấp lời giải thích hậu nghiệm về các hiện tượng thực nghiệm và hướng dẫn các lựa chọn thiết kế cấp cao. Phạm vi ngày càng tăng của học máy cũng đã dẫn đến việc kết hợp các lĩnh vực toán học mới, bao gồm tô-pô, đại số và hình học. Những lĩnh vực toán học này đã phát triển khả năng xử lý các mức độ trừu tượng và phức tạp cao, khiến chúng phù hợp để giải quyết các thách thức trong học sâu hiện đại. Các nhà nghiên cứu cho rằng toán học sẽ tiếp tục đóng vai trò quan trọng trong học máy, mặc dù trong một bối cảnh liên ngành và ứng dụng hơn.

8 Nhân viên Amazon đối mặt nguy cơ bị sa thải vì ủng hộ giới hạn data center

Amazon employees say they're facing termination for backing data center limits

The Verge AI Đọc bài viết →

Các kỹ sư phần mềm của Amazon, Patrick Schloesser, Darius Irani và Liesl Wigand, đã cáo buộc người sử dụng lao động của họ trả thù họ vì lên tiếng ủng hộ các quy định về trung tâm dữ liệu. Ba kỹ sư này đã làm chứng tại một phiên điều trần của Hội đồng Thành phố Seattle vào đầu tháng này, dẫn chiếu đến một luật thành phố cấm phân biệt đối xử việc làm về ngôn từ chính trị. Tuy nhiên, một tuần sau phiên điều trần và một ngày sau khi hội đồng thông qua một lệnh tạm ngừng xây dựng trung tâm dữ liệu, họ đã được gọi vào các cuộc họp với bộ phận Nhân sự của Amazon, nơi họ được thông báo rằng công ty đang điều tra họ và rằng hành động kỷ luật, bao gồm cả sa thải, là có thể. Các kỹ sư đã nộp đơn khiếu nại lên Văn phòng Quyền Dân sự Seattle, cáo buộc Amazon đã tham gia vào phân biệt đối xử việc làm bị cấm. Amazon đã bác bỏ cách mô tả tình huống, tuyên bố rằng nhân viên không có nguy cơ bị sa thải và rằng công ty không dung túng cho hành vi trả thù. Sự việc này xảy ra khi Seattle ban hành một lệnh tạm ngừng một năm đối với các trung tâm dữ liệu quy mô lớn, và các kỹ sư là thành viên của nhóm Amazon Employees for Climate Justice, một nhóm chuyên giải quyết khủng hoảng khí hậu.

TIPS & TRICKS CHO DEV

Tối ưu hóa GitHub Copilot

Vấn đề: Thiếu hiểu biết về code suggestions.

Cách làm: Sử dụng GitHub Copilot trong IDE, nhập lệnh `github.copilot.enable` để kích hoạt. Ví dụ, nhập `function greet` và Copilot sẽ đề xuất `function greet(name: string) { console.log(` Hello, ${name}! `); }`.

Đánh giá: Hiệu quả khi viết code nhanh, nhưng nên xem xét kỹ lưỡng đề xuất.

Sử dụng Windsurf

Vấn đề: Không thể tự động hóa code completion.

Cách làm: Cài đặt Windsurf, nhập lệnh `windsurf init` để cấu hình. Ví dụ, nhập `windsurf suggest` và cung cấp thông tin về dự án để nhận đề xuất.

Đánh giá: Hiệu quả khi cần tự động hóa code completion, nhưng yêu cầu cấu hình ban đầu.

Tích hợp Continue.dev

Vấn đề: Thiếu kiểm soát code chất lượng.

Cách làm: Tích hợp Continue.dev vào dự án, nhập lệnh `continue init` để bắt đầu. Ví dụ, nhập `continue analyze` và nhận phân tích về chất lượng code.

Đánh giá: Hiệu quả khi cần kiểm soát chất lượng code, nhưng yêu cầu cấu hình và tích hợp.

BÀI HỌC AI HÔM NAY CHO DEV

1. **Tối ưu chi phí & hiệu năng LLM**
2. Tối ưu chi phí và hiệu năng LLM giúp giảm thiểu chi phí vận hành và tăng tốc độ xử lý của mô hình AI. Điều này đặc biệt quan trọng khi triển khai ứng dụng AI trên quy mô lớn. Dev cần biết cách tối ưu hóa LLM để đảm bảo hiệu suất và tiết kiệm chi phí.
3. Ví dụ, sử dụng kỹ thuật fine-tuning và LoRA (Low-Rank Adaptation) có thể giúp giảm kích thước mô hình và tăng tốc độ xử lý. Code snippet minh họa: `model = transformers.AutoModelForSequenceClassification.from_pretrained('distilbert-base-uncased', num_labels=8)`
4. Tip: Sử dụng thư viện như Hugging Face Transformers để tối ưu hóa LLM và giảm thiểu chi phí vận hành. Tiếp theo, hãy thử nghiệm với các kỹ thuật tối ưu hóa khác như quantization và pruning để đạt được hiệu suất cao hơn.