

Relatório 1 - Vídeo: Fundamentos de Agentes de IA (I)

Isadora Stéfany Rezende Remigio Mesquita

Descrição da atividade

O objetivo dessa atividade foi fundamentar e consolidar o meu conhecimento a respeito do que é um Agente de IA. Não apenas um Agente em si, mas como são estruturados fluxos de trabalho que utilizam Agente ou Agentes de IA, os quais podem ser combinados na construção de uma solução robusta. Além disso, estudei a hierarquia (pilha abordada no vídeo 2) que permite que esses agentes e suas aplicações existam, deixando evidente como o desenvolvimento de protótipos e soluções inteiras com IA podem acontecer muito mais rápido que antes.

Vídeo 1: O que são Agentes de IA?

Neste vídeo, a apresentadora explica como os Agentes de IA evoluíram a partir do surgimento dos LLMs e o que definem eles. De forma concisa, são sistemas compostos que integram modelos de linguagem com componentes programáticos, permitindo resolver problemas complexos de forma mais eficiente. Um grande diferencial desses Agentes é a capacidade de raciocínio e a utilização de ferramentas externas, tornando-os cada vez mais autônomos e adaptáveis.

Como o próprio vídeo explica, os modelos por si só (modelos tradicionais) são limitados pelos dados de treinamento, o que os torna difíceis de adaptar, exige alta carga de dados de treinamento e esforço de desenvolvimento. Os Agentes de IA lidam com LLMs que possuem um conhecimento amplo, e com a capacidade de ativar outras ferramentas, tornam a sua adaptação a contextos diversos e treinamento muito mais fáceis, pois eles podem, por exemplo, realizar consultas na web.

Porém, isso só pode ocorrer devido à capacidade de aplicar conhecimentos de System Design. Ou seja, em um sistema há diversos componentes, ou seja, são inerentemente modulares. Dessa forma, posso ter tanto um modelo LLM ajustado quanto etapas programáticas que podem estar associadas a ele. Essas etapas podem ser: verificadores de output, realizar uma consulta e mapeá-la para aumentar as chances de que o retorno esteja correto. Enfim, o que podemos compreender bem é que antes tínhamos a parte programática e modelos generativos. Com a associação desses dois, especialmente através de Agentes, o desenvolvimento de novas soluções está cada vez mais acessível.

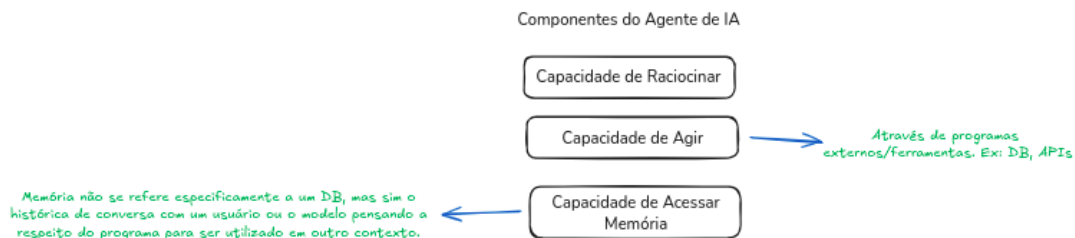


Figura 1: Componentes de um Agente de IA.

O ReACT é exatamente sobre o que um agente é capaz de fazer dados os seus componentes. Ou seja, o próprio nome indica ser Raciocínio + Ação.

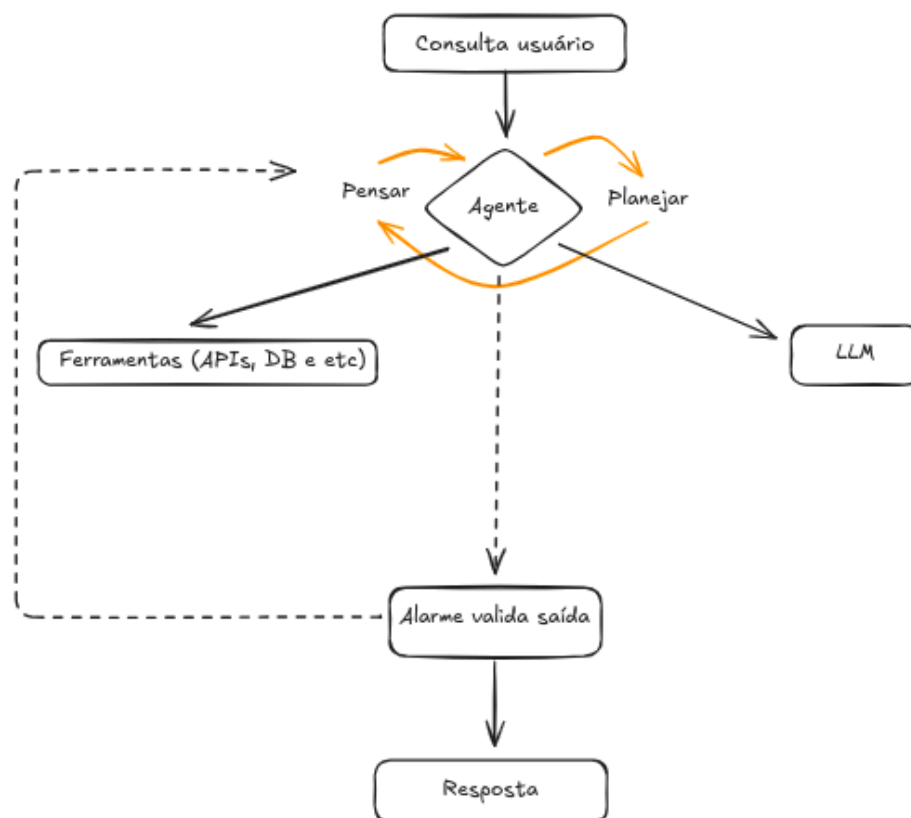


Figura 2: Fluxo de Decisão de um Agente de IA.

Vídeo 2: Andrew Ng Explores The Rise Of AI Agents And Agentic Reasoning | BUILD 2024 Keynote.

Nesse vídeo, Andrew fala sobre como é construída a pilha da IA, e ele aponta como a camada de orquestrados de Agentes é uma nova camada que está gerando um grande impacto nas aplicações, possibilitando a produção de dezenas de protótipos de uma aplicação de forma paralela e em um curtíssimo tempo, como uma semana.

Não se limitando a LLMs, Andrew também aborda sobre como modelos multimodais também podem ser utilizados nessas aplicações, e enfatiza como ele tem grandes expectativas em relação a modelos de imagem.

Além disso, foram discutidos também os tipos de raciocínio de um agente (reflexão com LLM, com as ferramentas que pode se relacionar, multiagentes e avaliar a solicitação do usuário para tomar decisão a respeito da melhor solução). Por fim, Andrew finaliza deixando suas 4 “apostas” sobre o que acredita que serão tendências, deixando clara a ascensão da engenharia de dados, uma vez que as outras etapas de um projeto também precisarão ganhar celeridade, dado o tempo atual de criação de um projeto de ML.

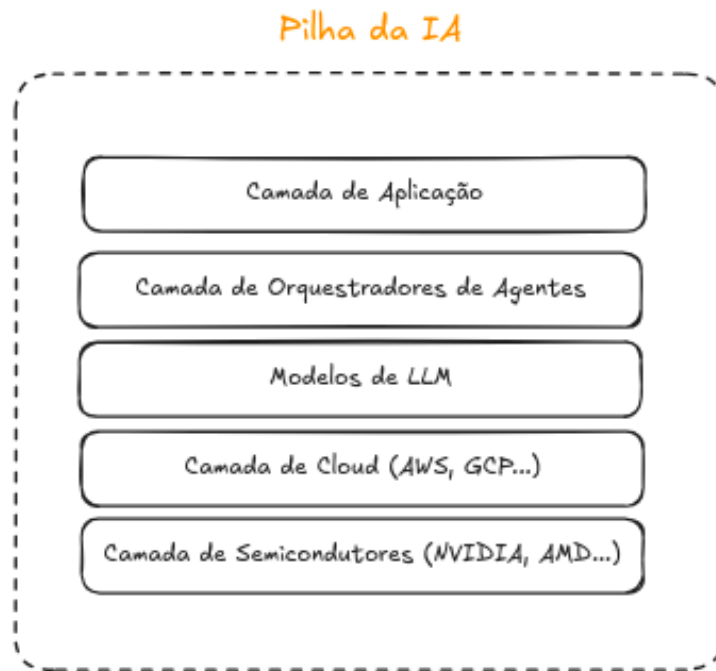


Figura 3: Pilha da IA discutida por Andrew.

Dificuldades

Não percebi dificuldades.

Conclusões

Na minha percepção ficou claro que o agente atua como um gerente de grande parte ou de quase todas as operações para um fim e/ou solicitação específica. Ele é capaz de administrar a conexão com DB, pesquisa web, analisar as solicitações e aprender com as solicitações. Além disso, foi uma novidade para mim agentes que lidam com dados além dos textuais.

Referências

<https://www.youtube.com/watch?v=F8NKhkZZWI>

<https://www.youtube.com/watch?v=KrRD7r7y7NY>

<https://mittechreview.com.br/futuro-agentes-ia-2025/>