

Projeto de Agentes de IA: Fundamentos e Aplicações

Objetivo da disciplina

Capacitar os alunos (graduação e pós-graduação) a projetar, implementar e avaliar sistemas de agentes de inteligência artificial baseados em modelos de linguagem de grande escala (LLMs). A disciplina oferece uma formação integrada que combina aspectos teóricos fundamentais com práticas de desenvolvimento, permitindo criar sistemas agênticos autônomos e multi-agentes para resolver problemas complexos em diversos domínios.

Ementa

Fundamentos de agentes de IA (AI Agents) e sistemas agênticos (Agentic AI). Modelos de linguagem de grande escala como base para agentes autônomos. Arquiteturas de agentes: percepção, raciocínio, planejamento e ação. Sistemas multi-agentes: colaboração, comunicação e coordenação. Técnicas de aprendizado por reforço para agentes. Uso de ferramentas e APIs externas. Memória e gestão de contexto em agentes. Avaliação e benchmarking de sistemas agênticos. Desafios éticos e de segurança. Aplicações práticas em ciência, engenharia e negócios.

Programa

1. Fundamentos de agentes de IA

Definição de agentes autônomos e sistemas agênticos. Diferenças entre IA generativa, agentes de IA baseados em LLM e sistemas agênticos completos, incluindo suas capacidades de autonomia e tomada de decisão.

2. Modelos de linguagem e agentes

Arquitetura e funcionamento de LLMs como módulo central de agentes. Como os LLMs realizam raciocínio, planejamento e tomada de decisões, incluindo técnicas de prompt engineering e estratégias de inferência.

3. Arquitetura de agentes baseados em LLM

Componentes fundamentais: perfil do agente, percepção, memória, planejamento e ação. Frameworks de implementação como ReAct, que integra raciocínio e ação de forma iterativa para resolução de problemas complexos.

4. Planejamento e raciocínio

Decomposição de tarefas complexas em sub-tarefas gerenciáveis. Técnicas de planejamento de curto e longo prazo, incluindo estratégias para lidar com incerteza e ajustar planos dinamicamente.

5. Uso de ferramentas e APIs

Integração de agentes com ferramentas externas, APIs e bancos de dados. Técnicas de function calling e recuperação aumentada por informação (RAG) para expandir as capacidades dos agentes além do conhecimento interno.

6. Memória e gestão de contexto

Estratégias para manutenção de memória de curto e longo prazo em agentes. Gerenciamento eficiente de contexto em conversas extensas e persistência de informações entre sessões.

7. Sistemas Multi-Agentes

Projeto de sistemas com múltiplos agentes especializados que colaboram. Estratégias de comunicação, coordenação e divisão de trabalho, incluindo padrões como orquestração centralizada e colaboração peer-to-peer.

8. Aprendizado por reforço multi-agente

Fundamentos de MARL (Multi-Agent Reinforcement Learning). Como agentes aprendem políticas ótimas em ambientes compartilhados através de interação e feedback.

9. Avaliação e benchmarking

Métricas e benchmarks para avaliar desempenho de agentes. Avaliação de qualidade de respostas, eficiência computacional, latência e capacidades específicas como uso de ferramentas e planejamento.

10. Desafios técnicos

Problemas como alucinações, rigidez de comportamento, comportamentos emergentes não desejados. Estratégias de mitigação incluindo loops de validação, calibração e modelagem causal.

11. Segurança e Ética

Considerações éticas no desenvolvimento de agentes autônomos. Questões de privacidade, alinhamento de valores, transparência e responsabilidade em sistemas agênticos.

12. Aplicações práticas

Casos de uso em descoberta científica, automação de processos, assistentes inteligentes. Discussão de sistemas reais implementados em domínios como saúde, finanças e robótica.

Avaliação

1. Avaliação teórica (40%)
2. Avaliação prática (60%)

Bibliografia

Livros

- Building Applications with AI Agents: Designing and Implementing Multiagent Systems, Michael Albada, 2025.
- The Agentic AI Bible: The Complete and Up-to-Date Guide to Design, Build, and Scale Goal-Driven, LLM-Powered Agents that Think, Execute and Evolve, Thomas R. Caldwell, 2025.
- AI Engineering: Building Applications with Foundation Models, Chip Huyen, 2024
- Artigos Científicos e Surveys