



**CAUÃ MARCOS DE OLIVEIRA SILVA  
JOÃO PEDRO NOGUEIRA LUCAS  
LUIZ VICTOR SORIANO DA CONCEIÇÃO  
MARDEM ARANTES DE CASTRO**

## **RELATÓRIO FINAL**

Projeto Prático de Sistemas Distribuídos

**LAVRAS - MG  
2024**

## Sumário:

|                                   |          |
|-----------------------------------|----------|
| <b>1 INTRODUÇÃO.....</b>          | <b>3</b> |
| <b>2 DESENVOLVIMENTO.....</b>     | <b>4</b> |
| 2.1 Agentes.....                  | 4        |
| 2.2 Backend.....                  | 4        |
| 2.3 Frontend.....                 | 4        |
| 2.4 Discussão e Resultados.....   | 4        |
| 2.4.1 Importância do Projeto..... | 5        |
| 2.4.2 Casos de Uso Reais.....     | 5        |
| <b>3 CONCLUSÃO.....</b>           | <b>6</b> |
| <b>4 REFERÊNCIAS.....</b>         | <b>7</b> |

## **1 INTRODUÇÃO**

Este projeto tem como objetivo o desenvolvimento de um sistema de chat baseado em modelos de linguagem (LLM) capaz de interagir com os usuários sobre conteúdos extraídos de arquivos CSV fornecidos. O sistema permite a extração e organização de informações de maneira eficiente, garantindo transparência, segurança e conformidade com as leis de proteção de dados pessoais. A escolha deste tema justifica-se pela necessidade de análise de dados em arquivos extensos, como grandes tabelas de dados, que são comuns em diversos setores, como agricultura, pesquisa científica e gestão de recursos naturais.

Na aplicação desenvolvida, o usuário pode interagir com agentes inteligentes sobre assuntos gerais e, em especial, sobre arquivos CSV. Essa funcionalidade torna o sistema útil em casos reais que demandam análise de tabelas, como a predição dos teores de areia, silte e argila em solos a partir de dados de elementos químicos (do Al ao Zr). Além disso, o sistema foi projetado para ser acessível a usuários não técnicos, como agricultores, que podem fazer perguntas em linguagem natural sem a necessidade de conhecimentos avançados em programação ou estatística.

## **2 DESENVOLVIMENTO**

O sistema foi desenvolvido com uma arquitetura modular composta por três principais componentes: agentes, backend e frontend. Cada componente desempenha um papel crucial no funcionamento do sistema, garantindo eficiência, segurança e usabilidade. Dessa forma, o código-fonte completo do projeto está disponível no repositório do GitHub: [github.com/LuizSoriano/TrabalhoSistemasDistribuïdos](https://github.com/LuizSoriano/TrabalhoSistemasDistribuïdos)

### **2.1 Agentes**

Os agentes são responsáveis pelo processamento das interações e pela lógica de negócios do chat. O agente principal utiliza um modelo de linguagem treinado para interpretar comandos do usuário e gerar respostas com base nos dados do arquivo CSV fornecido. A lógica de processamento envolve a extração de dados relevantes, a formatação adequada das respostas e a gestão do histórico de conversa para fornecer um diálogo coeso. Dois agentes foram implementados:

- Agente de Machine Learning (ML): Utiliza o modelo Random Forest para prever os teores de areia, silte e argila a partir de dados químicos do solo.
- Agente de Linguagem Natural (LLM): Baseado no modelo Ollama, permite que os usuários façam perguntas em linguagem natural sobre os dados do arquivo CSV.

### **2.2 Backend**

A API REST, implementada com Node.js, atua como intermediário entre o frontend e os agentes, gerenciando as requisições dos usuários e garantindo a segurança dos dados processados. O backend recebe os comandos do frontend e os encaminha ao agente responsável, retornando a resposta que será exibida ao usuário via chat.

### **2.3 Frontend**

A interface do usuário foi desenvolvida com HTML, CSS e JavaScript, proporcionando uma experiência interativa e intuitiva. O frontend permite que o usuário envie arquivos CSV, visualize os dados processados e interaja com o chatbot de maneira fluida, além de contar com um histórico de conversações para facilitar o acompanhamento das interações.

### **2.4 Discussão e Resultados**

A aplicação desenvolvida tem uma alta aplicabilidade no mundo real, especialmente em setores como agricultura, meio ambiente, pesquisa científica e gestão de recursos naturais. Sua importância reside na capacidade de transformar dados brutos (características do solo) em insights acionáveis, democratizando o acesso à análise técnica e à tomada de decisões baseadas em dados.

#### 2.4.1 Importância do Projeto

- **Democratização da Análise de Dados**
  - A integração com uma LLM permite que usuários não técnicos (ex.: agricultores) façam perguntas em linguagem natural, sem precisar de conhecimentos em programação ou estatística.
  - A interface amigável facilita o uso por profissionais de diferentes áreas, promovendo a inclusão digital.
- **Escalabilidade e Baixo Custo**
  - Ollama/Docker: Permitem deploy em nuvem ou servidores locais, escalando conforme a demanda.
  - Random Forest: Modelo leve e eficiente para previsões em tempo real, mesmo em hardware modesto.
  - Custo-Benefício: Substitui soluções caras (ex.: softwares proprietários de análise de solo) por uma alternativa acessível.
- **Inovação em Setores Tradicionais**
  - Agricultura 4.0: Alinha-se a tendências como IoT e big data no campo, modernizando práticas agrícolas.
  - Integração com Outras Tecnologias: Dados do solo podem ser combinados com imagens de satélite ou sensores IoT para análises mais robustas.

#### 2.4.2 Casos de Uso Reais

- **Cafeicultura no Brasil:** Uma cooperativa de café usa a aplicação para identificar solos com alto teor de silte, ideal para variedades de café arábica, otimizando o plantio.
- **Combate à Desertificação na África:** ONGs analisam tendências de degradação do solo em regiões semiáridas e planejam ações de reflorestamento.

### **3 CONCLUSÃO**

O desenvolvimento deste sistema proporcionou um amplo aprendizado sobre a importância da proteção de dados pessoais e dos desafios envolvidos na implementação de soluções seguras. A análise de riscos realizada permitiu identificar potenciais vulnerabilidades e propor soluções eficazes para mitigação desses problemas.

O projeto demonstrou a viabilidade de construir um sistema de chat interativo baseado em LLMs, garantindo a conformidade com normas de proteção de dados. Além disso, a aplicação de metodologias consagradas na segurança da informação reforça a relevância e confiabilidade da solução desenvolvida.

Por fim, espera-se que os conhecimentos adquiridos neste projeto possam ser aplicados em futuras implementações, contribuindo para a construção de sistemas cada vez mais seguros e eficientes. A aplicação desenvolvida não apenas atende às necessidades atuais de análise de dados, mas também abre caminho para inovações em setores tradicionais, promovendo a democratização do acesso à tecnologia e à informação.

#### 4 REFERÊNCIAS

SHARIFMOUSAVI, Mahdi; KAYVANFAR, Vahid; BALDACCI, Roberto. **Distributed Artificial Intelligence Application in Agri-food Supply Chains 4.0**. Procedia Computer Science, v. 232, p. 211-220, 2024. DOI: 10.1016/j.procs.2024.01.021.

Rohit Nishant, Mike Kennedy, and Jacqueline Corbett. **Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda**, International Journal of Information Management, 53, 102104, 2020.