CAUÃ MARCOS DE OLIVEIRA SILVA

JOÃO PEDRO NOGUEIRA LUCAS

LUIZ VICTOR SORIANO DA CONCEIÇÃO

MARDEM ARANTES DE CASTRO

RELATÓRIO DE IMPACTO À PROTEÇÃO DE DADOS

Projeto Prático de Sistemas Distribuídos

LAVRAS - MG

2024

Sumário:

1 INTRODUÇÃO

O presente relatório tem como objetivo avaliar os riscos à proteção de dados associados à implementação de um sistema distribuído para predição dos teores de areia, silte e argila em solos a partir de dados de elementos químicos (do Al ao Zr). O sistema foi desenvolvido como parte do projeto prático da disciplina de Sistemas Distribuídos (GCC129) do curso de Ciência da Computação e utiliza uma arquitetura baseada em agentes especializados: um agente de Machine Learning (ML) para predições e um agente de Linguagem Natural (LLM) para interação com os usuários.

Apesar de o sistema não lidar diretamente com dados pessoais, ele processa informações técnicas sensíveis, como dados químicos de solos, que podem ter valor estratégico para usuários, como agricultores, pesquisadores ou empresas do agronegócio. Portanto, a proteção desses dados contra vazamentos, adulterações ou acessos não autorizados é fundamental para garantir a confidencialidade, integração e disponibilidade do sistema.

Este relatório utiliza a metodologia STRIDE (Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, Elevation of Privilege) e a modelagem de ameaças de Torr (2005) para identificar e avaliar os principais riscos associados ao sistema. Além disso, são propostas medidas de mitigação para cada risco identificado, visando assegurar a segurança e a privacidade dos dados processados.

2 DESENVOLVIMENTO

2.1 Metodologia

O sistema foi desenvolvido com uma arquitetura modular composta por três principais componentes: agentes, backend(API REST) e frontend. A avaliação de riscos foi conduzida com base na metodologia STRIDE, que categoriza as ameaças em seis tipos principais: Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service e Elevation of Privilege. Além disso, a modelagem de ameaças de Torr (2005) foi utilizada para identificar vulnerabilidades específicas em cada componente do sistema.

2.2 Identificação e Avaliação dos Riscos

2.2.1 Diagrama de Fluxo de Dados (DFD)

Antes de iniciar a avaliação dos riscos, foi desenvolvido um diagrama de fluxo para mapear o funcionamento do sistema e identificar os pontos críticos onde os dados são processados, armazenados e transmitidos. Esse diagrama foi essencial para compreender o ciclo de vida dos dados dentro do sistema e para identificar possíveis vulnerabilidades em cada etapa. O fluxo começa com o usuário interagindo com a interface web, onde ele pode selecionar o agente desejado (Fazer predições ou perguntar ao LLM), enviar um arquivo CSV e digitar uma mensagem (prompt). O frontend, responsável pela interface, valida os dados inseridos e os envia ao backend por meio de requisições Fetch API.

No backend, os dados são recebidos e encaminhados para o agente correspondente. Se o usuário selecionou o agente ML, o backend envia o arquivo CSV para processamento, onde o agente ML aplica técnicas de pré-processamento, como transformações logarítmicas e normalização, e utiliza o modelo Random Forest para gerar as predições de areia, silte e argila. Essas predições são retornadas ao backend em formato JSON. Por outro lado, se o usuário selecionou o agente LLM, o backend envia o prompt e o caminho do arquivo CSV para o agente LLM, que utiliza o modelo Ollama para processar a pergunta e gerar uma resposta com base nos dados do arquivo. A resposta é retornada ao backend em formato de texto.

O diagrama de fluxo também destaca os pontos de comunicação entre os componentes, como a transmissão de dados entre o frontend e o backend, e entre o backend e os agentes especializados. Esses pontos foram analisados para identificar possíveis

vulnerabilidades, como a exposição de dados, o acesso não autorizado a arquivos temporários e a manipulação de dados durante o processamento.

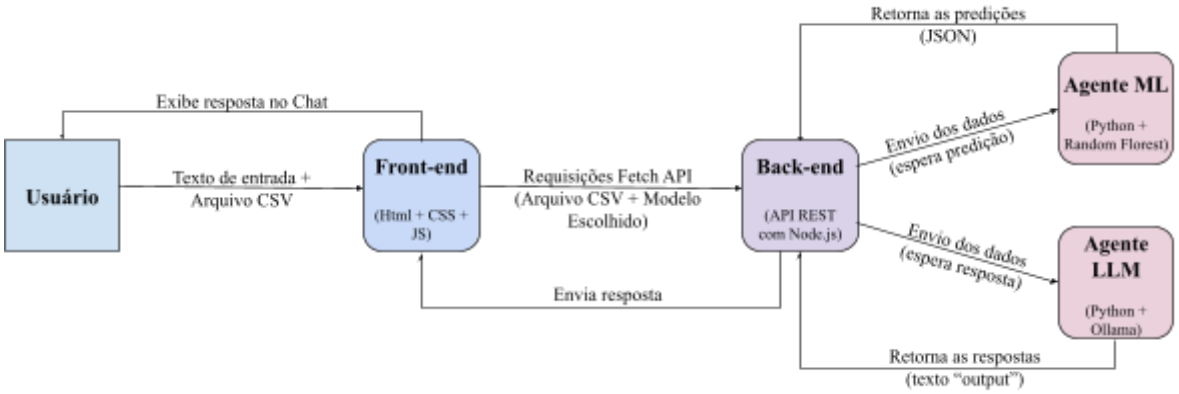

Figura 1: Diagrama de Fluxo de Dados (DFD)

2.2.2 Risco (Upload de CSV) - Exposição de Dados Técnicos Sensíveis

- Descrição: Arquivos CSV podem conter dados técnicos estratégicos (ex.: níveis de pH, teor de nutrientes) que, se expostos, podem ser usados por concorrentes para inferir práticas agrícolas específicas.
- Categoria STRIDE: Information Disclosure
- Probabilidade: Média
- Impacto: Médio (vantagem competitiva indevida)

2.2.2 Risco (Ollama) - Respostas com Padrões Estatísticos Sensíveis

- Descrição: A LLM pode gerar respostas que revelam padrões estatísticos críticos (ex.: "O maior teor de areia do conjunto de dados é 85%"), permitindo que terceiros reconstruam parcialmente o dataset original.
- Categoria STRIDE: Information Disclosure
- Probabilidade: Baixa
- Impacto: Médio (exposição de propriedade intelectual)

2.2.3 Risco (Modelo Random Forest) - Adulteração de Previsões

- Descrição: Substituição do modelo treinado por um malicioso, gerando previsões incorretas de Areia/Silte/Argila, o que pode levar a decisões agrícolas equivocadas.
- Categoria STRIDE: Tampering
- Probabilidade: Baixa
- Impacto: Alto (prejuízos operacionais).

2.2.4 Risco (NodeJS Backend) - Upload de CSV com Metadados Maliciosos

- Descrição: Arquivos CSV com metadados manipulados (ex.: nomes de colunas maliciosos) podem explorar vulnerabilidades no parser de dados.
- Categoria STRIDE: Spoofing
- Probabilidade: Baixa
- Impacto: Médio (interrupção do serviço)

2.2.5 Risco (Docker) - Acesso a Modelos ou Dados de Treinamento

- Descrição: Containers mal configurados podem permitir acesso não autorizado ao modelo treinado ou a dados históricos de solo usados no treinamento.
- Categoria STRIDE: Elevation of Privilege
- Probabilidade: Baixa
- Impacto: Alto (vazamento de propriedade intelectual)

2.2.6 Risco (Armazenamento) - Acesso Não Autorizado entre Usuários

- Descrição: Falha no isolamento de dados pode permitir que um usuário acesse CSVs de outro (ex.: dados de solo de um concorrente).
- Categoria STRIDE: Information Disclosure
- Probabilidade: Baixa
- Impacto: Alto (competitividade comprometida)

2.3 Gráfico de análise de riscos

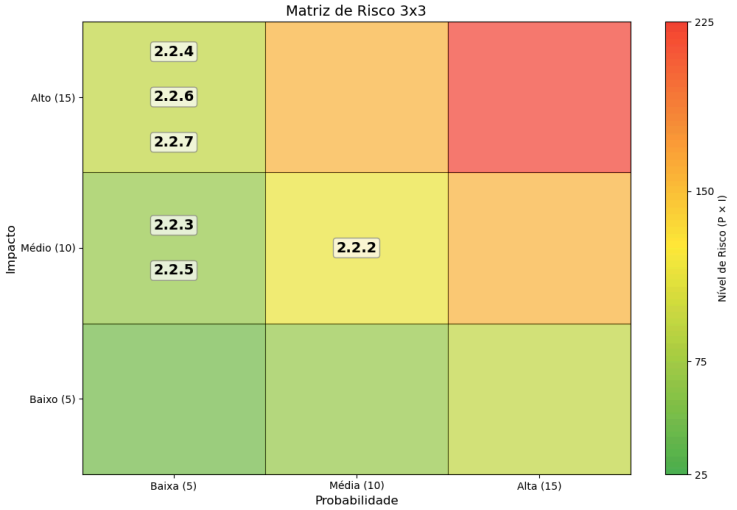

Figura 2: Gráfico de análise de riscos

2.4 Medidas para Tratar os Riscos

2.4.1 Medidas Gerais

- Minimização de Dados:
    - Exclusão automática de CSVs após processamento (ex.: 24 horas).
    - Restrição de colunas permitidas nos CSVs (ex.: apenas "pH", "Areia Total", "Silte", "Argila").
- Criptografia:
    - Dados em repouso criptografados com AES-256.
    - Uso de HTTPS obrigatório para todas as comunicações.

2.4.2 Medidas Específicas por Componente

Componente	**Risco**	**Medidas**
Upload de CSV	Exposição de dados técnicos	- Validação estrita de schema no upload (ex.: rejeitar CSVs com colunas fora do escopo esperado). - Remoção de metadados não essenciais (ex.: datas, comentários).
Ollama (LLM)	Revelação de padrões estatísticos	- Limitar respostas a estatísticas agregadas (ex.: médias, desvios padrão) sem mencionar extremos. - Implementar "differential privacy" em respostas da LLM.
Modelo Random Forest	Adulteração do modelo	- Assinatura digital do modelo (SHA-256) e verificação antes da inferência. - Ambiente de execução isolado (ex.: container Docker sem permissão de escrita).
NodeJS Backend	CSV malicioso	- Uso de bibliotecas seguras para parsing (ex.: csv-parser com validação de tipos). - Sandboxing de processos que manipulam CSVs.

Docker	Acesso a modelos/dados	- Configuração de containers em modo "read-only". - Uso de Docker Secrets para armazenar chaves de criptografia.
Armazenamento	Acesso cruzado entre usuários	- Isolamento físico ou lógico de dados por usuário (ex.: diretórios separados com permissões específicas). - Uso de IAM (Identity and Access Management) para controle granular.

2.4.3 Medidas Adicionais

- Monitoramento Proativo:
    - Alertas para atividades suspeitas (ex.: múltiplas tentativas de upload de CSVs fora do padrão).
    - Auditoria semanal de logs de acesso aos dados.

- Testes de Segurança Específicos:
    - Pentests focados em tentativas de reconstrução de datasets a partir de respostas da LLM.
    - Validação de integridade do modelo Random Forest após atualizações.

- Conformidade e Transparência:
    - Política clara de retenção de dados (ex.: CSVs são excluídos após processamento).
    - Relatório de segurança acessível aos usuários, detalhando medidas de proteção.

3 CONCLUSÃO

A análise de riscos realizada neste Relatório de Impacto à Proteção de Dados permitiu identificar ameaças significativas associadas ao sistema distribuído desenvolvido para predição dos teores de areia, silte e argila em solos. A metodologia STRIDE possibilitou uma avaliação detalhada das vulnerabilidades do sistema, abrangendo aspectos como spoofing, tampering, repudiation, information disclosure, denial of service e elevation of privilege. A implementação de medidas de segurança é essencial para garantir a proteção dos dados técnicos processados e a confiabilidade da solução.

Os riscos principais concentram-se na integridade das previsões do modelo, vazamento de dados técnicos estratégicos e ataques de negação de serviço. As medidas propostas garantem:

- **Confidencialidade:** Restrição de acesso e criptografia de dados técnicos.
- **Integridade:** Verificação de modelos e sanitização de entradas.
- **Disponibilidade:** Isolamento de containers e políticas de retenção curtas.

A aplicação alinha-se ao princípio de minimização de dados, processando apenas informações estritamente necessárias (ex.: atributos químicos do solo) e descartando-as após uso. Apesar da ausência de dados pessoais clássicos nos CSVs usados na aplicação, a proteção de propriedade intelectual e a prevenção de adulterações mantêm-se críticas para a confiabilidade do sistema.

4 REFERÊNCIAS

HOWARD, M.; WHITTAKER, J. **Demystifying the Threat-Modeling Process**. Disponível em: <https://www.ida.liu.se/~TDDC90/literature/papers/torr.pdf>. Acesso em: 11 de fev de 2025.