

Nome do Grupo: I9

Nome	E-mail
Brenik (responsável)	brenik.oliveira@hotmail.com
Bruno	bruno.lucena@cnh.com
Wei	wei.lai@cnhind.com
Bruno (desistente)	brunnobts@gmail.com
Wilmara (desistente)	wilmaragelape@yahoo.com.br
Tales (desistente)	talesnazareno@gmail.com
Gustavo(desistente)	govaty@gmail.com
Vanderleia (desistente)	vangutendorfer@gmail.com

Síntese de aprendizado:

O grupo multidisciplinar foi eficaz na concepção do projeto, mas perfis distintos seriam melhores na execução. Enfrentamos os mesmos desafios das empresas: boas ideias e motivação inicial, porém falta de estrutura para executar. O curso serviu para avaliarmos nosso grau de maturidade em projetos avançados de IA sem apoio externo. Temos força em ideação e mapeamento de valor, mas ainda temos longa jornada na parte de desenvolvimento. Optamos por usar apenas ferramentas gratuitas para testar novos recursos. A rotatividade da equipe também fez parte do experimento, trazendo aprendizado sobre retomada de atividades e redução de impactos.

No que diz respeito ao desenvolvimento, reconhecemos que este ainda representa um desafio para nossa equipe. No entanto, a evolução tecnológica dos últimos anos, especialmente no que se refere à geração e correção automática de código por LLMs, permitiu que conseguíssemos entregar um MVP. Em um passado recente, alcançar esse resultado em tão pouco tempo e com o atual perfil do time teria sido consideravelmente mais difícil.

O aprendizado durante o curso proporcionou uma visão abrangente sobre a evolução histórica da Inteligência Artificial, destacando o momento de democratização trazido pelas LLMs, a compreensão de sua arquitetura e funcionamento, o impacto do conceito de Retrieval-Augmented Generation (RAG) na superação das limitações dos modelos de linguagem, e, por fim, a experiência transformadora de desenvolver agentes autônomos capazes de planejar e executar ações, representando a próxima fronteira da IA aplicada.

Com o conhecimento adquirido ao longo do curso, estamos aptos a prototipar soluções e aprimorar a definição de requisitos para solicitar o desenvolvimento de projetos junto a empresas especializadas.

O curso foi utilizado para capacitar nosso líder mais jovem, proporcionando um exercício prático de aprimoramento em gestão do tempo, planejamento estratégico, execução e tomada de decisão. Adicionalmente, contribuiu para o desenvolvimento das habilidades de comunicação. O curso foi um ambiente seguro para o nosso líder experimentar os desafios desse papel e treinar as habilidades de liderança.

Proteção Automatizada para Pequenas e Médias Empresas

Descrição do Tema:

Conf.IA visa criar um agente inteligente capaz de processar e comparar dados de Notas Fiscais Eletrônicas (NFe) e seus respectivos DANFEs, ambos em formato digital. A conciliação se dará pela análise de campos-chave presentes em ambos os documentos, garantindo a conformidade e identificando possíveis erros ou fraudes.

Objetivos:

- Automatizar a conciliação de NFe e DANFE digitais.
- Reduzir o tempo gasto em processos manuais de verificação.
- Minimizar erros humanos na identificação de divergências.
- Fornece relatórios claros e acionáveis sobre as inconformidades encontradas.
- Melhorar a integridade dos dados fiscais da empresa.

Público-Alvo

Pequenas e Médias Empresas (PMEs) que:

- Processam entre 100 a 10.000 notas fiscais mensalmente
- Possuem equipes de auditoria limitadas ou terceirizadas
- Buscam reduzir custos operacionais com controles internos
- Necessitam de conformidade regulatória
- Querem proteger-se contra fraudes internas e externas

Segmentos Prioritários:

- Varejo e E-commerce
- Distribuidoras e Atacadistas
- Prestadores de Serviços
- Indústrias de Pequeno e Médio Porte

Justificativa do Tema

Por que é Importante?

Impacto Financeiro das Fraudes

- PMEs perdem em média 5% de sua receita anual devido a fraudes
- 42% das pequenas empresas sofrem fraudes internas
- Tempo médio para detectar fraudes: 14 meses sem automação

Limitações Atuais das PMEs

- Recursos Limitados: Equipes pequenas de auditoria

- Processos Manuais: Verificação manual é lenta e propensa a erros
- Falta de Expertise: Dificuldade em identificar padrões complexos
- Alto Custo: Soluções enterprise são inacessíveis

Valor Agregado ao Público-Alvo

Redução de Custos

- Diminuição de até 80% no tempo de análise
- Redução de perdas por fraudes em até 60%
- Otimização da equipe de auditoria

Eficiência Operacional

- Detecção proativa vs. reativa
- Priorização inteligente de casos suspeitos
- Automação de processos repetitivos

Conformidade e Segurança

- Atendimento às exigências regulatórias
- Criação de trilha de auditoria digital
- Proteção da reputação empresarial

Proposta de Desenvolvimento

Escopo Técnico

1. Fontes de Dados

- Notas Fiscais Eletrônicas (NFe): Arquivos XML são o formato preferencial. Alternativamente, dados podem ser fornecidos em CSV ou outros formatos tabulares, desde que contenham os campos essenciais.
- DANFEs: Espera-se que os DANFEs estejam em formato digital (por exemplo, PDF que possa ser parsado para extração de texto, ou dados já estruturados em CSV/XML se disponíveis por algum sistema). Não haverá OCR envolvido. A extração de dados do DANFE será baseada em parsing do texto contido no PDF (se for o caso) ou diretamente de formatos estruturados.

2. Módulos Principais

- Módulo de Ingestão de Dados:
 - Leitores de Arquivo: Responsável por ler e parsear os diferentes formatos de arquivo (XML para NFe, e possivelmente PDF para extração de texto do DANFE, ou CSV/JSON se o DANFE já vier estruturado).
 - Validação de Esquema: Garantir que os arquivos de entrada estejam em conformidade com os esquemas esperados (e.g., esquema XML da NFe).
- Módulo de Pré-processamento:
 - Extração de Campos-Chave: Identificar e extrair os campos relevantes para a conciliação de ambos os documentos (ver Seção 3.3).

- Normalização de Dados: Padronizar formatos (e.g., datas, valores numéricos, CNPJ/CPF) para garantir comparações precisas.
- Módulo de Conciliação:
 - Algoritmo de Comparação: Implementar lógica para comparar os campos-chave extraídos de NFe e DANFE.
 - Regras de Negócio: Definir as regras de conciliação, por exemplo, qual nível de tolerância para pequenas diferenças de valores, ou quais campos são mandatórios para uma conciliação "positiva".
 - Identificação de Discrepâncias: Marcar e categorizar as diferenças encontradas.
- Módulo de Relatórios e Alertas:
 - Geração de Relatórios: Apresentar as conciliações e, principalmente, as divergências encontradas de forma clara (e.g., CSV, Excel, PDF).
 - Sistema de Alerta: Notificar usuários ou sistemas externos sobre divergências críticas (e.g., e-mail, integração com fila de mensagens).

3. Campos para Conciliação (Exemplos)

Campo	NFe (XML)	DANFE (Digital)
Chave de Acesso	<infNFe Id="NFe...	Campo específico no DANFE (geralmente em destaque)
Número da NFe	<nNF>	Número da Nota Fiscal
Série da NFe	<serie>	Série da Nota Fiscal
Data de Emissão	<dhEmi>	Data de Emissão
CNPJ/CPF do Emitente	<CNPJ> / <CPF> do <emit>	CNPJ/CPF do Emitente
Razão Social do Emitente	<xNome> do <emit>	Razão Social do Emitente
CNPJ/CPF do Destinatário	<CNPJ> / <CPF> do <dest>	CNPJ/CPF do Destinatário
Razão Social do Destinatário	<xNome> do <dest>	Razão Social do Destinatário
Valor Total da NFe	<vNF>	Valor Total da Nota
Valor Total do ICMS	<vICMS>	Valor do ICMS (se aplicável e visível no DANFE)
Descrição de Produtos/Serviços	<xProd> em <det>	Descrição resumida de itens (se disponível e relevante)
Quantidade de Produtos/Serviços	<qCom> em <det>	Quantidade de itens (se disponível e relevante)
Unidade Comercial	<uCom> em <det>	Unidade de Medida (se disponível e relevante)
Valor Unitário Comercial	<vUnCom> em <det>	Valor Unitário (se disponível e relevante)

Nota: A seleção exata dos campos dependerá da criticidade e dos requisitos específicos da conciliação.

4. Tecnologias Sugeridas

4.1 Core code

1. Linguagem de Programação: Python é altamente recomendado devido à sua vasta biblioteca para manipulação de dados, parsing de XML/JSON/CSV, e capacidade de processamento de texto.
2. Processamento de XML: Bibliotecas como lxml ou ElementTree para parsear as NFes.
3. Processamento de PDF (para DANFE): Bibliotecas como PyPDF2, pdfminer.six ou fitz (PyMuPDF) para extrair texto de PDFs. Reiterando: sem uso de OCR. A extração será do texto "legível por máquina" dentro do PDF.
4. Manipulação de Dados: Pandas para estruturar, limpar e comparar os dados de forma eficiente.
 1. Banco de Dados (Opcional, mas recomendado para histórico e persistência):
 2. SQL: PostgreSQL ou MySQL para armazenar dados de NFe, DANFE e resultados da conciliação.
 3. NoSQL: MongoDB ou outro para armazenamento de documentos, se a flexibilidade for uma prioridade.
5. Orquestração/Automação: Ferramentas como Apache Airflow ou scripts agendados (e.g., cron jobs em Linux, Task Scheduler em Windows) para gerenciar o fluxo de processamento.
6. Versionamento de Código: Git/GitHub/GitLab.

4.2 Front End

O frontend desta aplicação foi desenvolvido utilizando **React** como biblioteca principal, aproveitando hooks modernos como useState, useRef e useEffect para gerenciar estado e ciclo de vida dos componentes de forma eficiente. A interface visual foi construída com **Tailwind CSS** através de classes utilitárias, criando um design moderno com gradientes escuros (slate e purple) e efeitos de glassmorphism. Para as visualizações de dados, foi integrada a biblioteca **Recharts**, que oferece componentes declarativos e responsivos para gráficos de barras, linhas, pizza, dispersão e área, todos com tooltips interativos e animações suaves. Os ícones são fornecidos pela biblioteca **Lucide React**, garantindo uma iconografia consistente e escalável. No backend de processamento, a aplicação utiliza **MathJS** para cálculos matemáticos avançados (média, mediana, correlação, desvio padrão e variância), **Lodash** para manipulação eficiente de arrays e objetos (agregações, agrupamentos e operações funcionais), e **SheetJS (XLSX)** para parsing de arquivos Excel, embora o código também suporte leitura nativa de CSV e JSON através da API FileReader do navegador. Toda a lógica de análise de dados é executada no lado do cliente, sem necessidade de servidor, proporcionando uma experiência rápida e privada para o usuário.

5. Framework Escolhido

LangChain é um framework open-source projetado para facilitar o desenvolvimento de aplicações baseadas em modelos de linguagem (LLMs), como agentes, chatbots, RAG (retrieval-augmented generation), entre outros.

Pontos Positivos:

- **Alta Flexibilidade para IA Generativa:** Suporte nativo à criação de agentes inteligentes, uso de ferramentas externas, memória conversacional e RAG.
- **Ampla Comunidade e Ecossistema:** Forte crescimento no ecossistema open-source, com contribuições regulares e suporte da comunidade.
- **Integração Avançada com APIs e Bases de Dados:** Permite criação de pipelines sofisticados que combinam IA, dados estruturados e não estruturados.
- **Escalável para Casos Avançados de IA:** Ideal para aplicações em larga escala com múltiplos componentes inteligentes.

Pontos Negativos:

- **Requer Conhecimento de Programação:** Totalmente voltado para desenvolvedores (Python/JS); não é adequado para times sem expertise técnica.
- **Curva de Aprendizagem Elevada:** A modelagem de agentes, memórias e cadeias exige conhecimento de conceitos avançados de IA.
- **Infraestrutura Necessária:** A integração com vetores, bancos de dados, APIs e hospedagem exige investimento em arquitetura técnica robusta.

6. Próximos Passos de Planejamento

1. Levantamento Detalhado de Requisitos:
 - Definir com precisão todos os campos a serem conciliados.
 - Estabelecer as regras de negócio para a conciliação (tolerâncias, campos obrigatórios, etc.).
 - Definir os formatos de saída dos relatórios e o sistema de alerta preferencial.
2. Coleta de Amostras de Dados: Obter um conjunto representativo de NFes (XML) e DANFEs digitais para testes.
3. Desenho Detalhado da Solução:
 - Modelagem de dados para o banco de dados (se aplicável).
 - Definição das APIs ou interfaces entre os módulos.
4. Seleção e Configuração de Ferramentas: Instalar e configurar o ambiente de desenvolvimento.
5. Desenvolvimento Iterativo: Construir e testar os módulos em fases.
6. Testes e Validação: Realizar testes unitários, de integração e de aceitação do usuário.

Modelo de Negócio

SaaS por Faturamento

- Starter: Até R\$ 500K faturamento/mês - R\$ 299/mês
- Professional: Até R\$ 2M faturamento/mês - R\$ 799/mês
- Enterprise: Acima R\$ 2M faturamento/mês - R\$ 1.999/mês

Implementação

- Setup inicial gratuito
- Treinamento da equipe incluso
- Suporte técnico 24/7
- Integrações personalizadas (taxas adicionais)

Métricas de Sucesso

KPIs Técnicos

- Taxa de detecção de fraudes: >85%
- Falsos positivos: <10%
- Tempo de processamento: <30 segundos por nota
- Disponibilidade do sistema: >99.5%

KPIs de Negócio

- Redução de perdas por fraude: >50%
- ROI do cliente: >300% no primeiro ano
- Tempo de implementação: <30 dias
- Satisfação do cliente: >4.5/5.0

Milestones

1. Validação do MVP com 5-10 empresas piloto
2. Captação de investimento para desenvolvimento completo
3. Contratação da equipe técnica especializada
4. Parcerias estratégicas com contadores e auditorias
5. Lançamento comercial após 9 meses de desenvolvimento

"Transformando a detecção de fraudes de reativa para proativa, protegendo o crescimento das PMEs brasileiras."

Agentes Autônomos de IA

A solução agora incorpora 4 agentes especializados:

- Agente Detector: Monitora transações 24/7
- Agente Analista: Correlaciona padrões complexos
- Agente Investigador: Aprofunda casos suspeitos
- Agente Comunicador: Gera alertas contextualizados

Investimento Total: R\$ 1.600.000

- Fase 1: R\$ 450.000 (Fundação - 3 meses)
- Fase 2: R\$ 600.000 (IA/ML - 3 meses)
- Fase 3: R\$ 550.000 (Otimização - 3 meses)
- Operacional: R\$ 1.170.000/ano

Equipe Suporte 24/7: R\$ 45.000/mês

- 3 Analistas Técnicos: R\$ 18.000
- 2 Especialistas IA: R\$ 24.000
- 1 Coordenador: R\$ 9.000

Proposta de Investimento Anjo — Plataforma Inteligente de Detecção de Fraudes em Notas Fiscais O

Problema

Empresas brasileiras perdem bilhões anualmente com fraudes fiscais — muitas vezes silenciosas e difíceis de detectar com métodos tradicionais. Essas fraudes geram não apenas prejuízos financeiros, mas também riscos legais e reputacionais.

Nossa Solução

Uma plataforma SaaS de **detecção de fraudes baseada em IA e Machine Learning** que analisa, em tempo real, padrões anômalos em milhares de notas fiscais eletrônicas emitidas por empresas. Automatizamos a auditoria, aumentamos a precisão e reduzimos drasticamente o risco fiscal.

Oportunidade de Investimento

Estamos buscando R\$ 1.600.000 em capital semente para alavancar as três fases do desenvolvimento da plataforma em 9 meses:

- Fase 1 – Fundação (R\$ 450.000 / 3 meses)
Estruturação técnica, arquitetura da solução, parcerias com fontes de dados e protótipo funcional.
- Fase 2 – Inteligência Artificial (R\$ 600.000 / 3 meses)
Desenvolvimento de modelos de IA/ML para detectar padrões suspeitos, integração com ERP e testagem com dados reais.
- Fase 3 – Otimização (R\$ 550.000 / 3 meses)
Refinamento de algoritmos, dashboard inteligente, automações de alerta e preparo para escalar.

Time Operacional e Suporte 24/7

Custo mensal: R\$ 45.000

- 3 Analistas Técnicos
- 2 Especialistas em IA
- 1 Coordenador de Operações

Projeções Financeiras

- **Break-even com 155 clientes:**
R\$ 128.845/mês em receita | 12 meses
- **Faturamento saudável com 330 clientes:**
R\$ 275.670/mês | 18 meses | **ROI estimado: 24–30 meses**
- **Escala otimizada com 550 clientes:** R\$ 459.450/mês | 24 meses

Por que Investir Agora?

- Mercado crescente com forte demanda por soluções fiscais inteligentes
- Modelo SaaS com **alto potencial de escalabilidade e receita recorrente**
- Solução defensável com base em **dados proprietários e algoritmos exclusivos**
- Equipe técnica sólida e pronta para execução
- Roadmap validado com parceiros pilotos já em negociação

Proposta ao Investidor Anjo

- **Investimento:** R\$ 1.600.000
- **Participação sugerida:** a negociar, com base em valuation pré-money (ex: 15–20%)
- **Opção de saída:** rodada Série A (18–24 meses), fusão ou aquisição por ERP/fisco/fintechs •

Potencial de múltiplos (x6–x10) em até 3 anos

Viabilidade Financeira

Para Break-even (155 clientes):

- Faturamento: R\$ 128.845/mês
- Timeline: 12 meses

Para Faturamento Saudável (330 clientes):

- Faturamento: R\$ 275.670/mês
- Timeline: 18 meses
- ROI: 24-30 meses

Escala Otimizada (550 clientes):

- Faturamento: R\$ 459.450/mês
- Timeline: 24 meses

A proposta agora apresenta um modelo completo com agentes autônomos, custos detalhados e projeções financeiras realistas para PMEs brasileiras.

