

Relatório de entrega final: Analisador e Classificador Inteligente de Documentos Fiscais

Grupo e integrante: Anna Luiza
Nome do grupo: Anna_Milene_Jonatas

1. Descrição do Tema Escolhido

O projeto desenvolvido consiste em um Agente de Inteligência Artificial para Análise e Classificação de Documentos Fiscais (NF-e). O objetivo central foi criar uma solução robusta e adaptável capaz de processar documentos fiscais (XML e PDF), extrair dados cruciais e, fundamentalmente, classificar e categorizar essas operações com base no ramo de atividade da empresa.

A principal inovação reside na implementação de uma Arquitetura de Agentes Especializados. Em vez de uma lógica de classificação monolítica, o sistema utiliza um Agente Orquestrador que delega a análise fiscal a agentes setoriais específicos (como Agronegócio, Setor Automotivo e Indústria). Essa modularidade permite que as regras de negócio, alertas e recomendações sejam altamente customizadas para as particularidades tributárias de cada setor.

2. Público-Alvo

A solução é direcionada a profissionais e empresas que lidam com um alto volume de documentos fiscais e que necessitam de precisão e conformidade na gestão tributária e contábil. O público-alvo primário inclui:

Categoria	Perfil	Benefício Principal
Analistas Fiscais e Contadores	Profissionais responsáveis pela apuração de impostos, escrituração e conformidade legal.	Redução drástica do tempo gasto com classificação manual e conferência de regras específicas.
Gestores Financeiros e Controladoria	Responsáveis pela gestão de custos, centros de custo e planejamento tributário.	Geração automática de insumos para cálculo de custos de produção e alocação precisa de despesas.
Pequenas e Médias Empresas (PMEs)	Empresas que buscam automatizar processos fiscais sem a necessidade de grandes sistemas ERP.	Acesso a inteligência fiscal especializada e organização documental eficiente.

3. Justificativa do Tema Escolhido

A complexidade do sistema tributário brasileiro, aliada à constante mudança na legislação, torna a gestão fiscal um desafio oneroso e de alto risco para as empresas. O tema escolhido se justifica pela necessidade de automação inteligente e especialização setorial, agregando os seguintes valores:

- Conformidade e Redução de Riscos: Ao aplicar regras fiscais customizadas por ramo de atividade (ex: monitoramento de IPI para Indústria, FUNRURAL para Agronegócio), o agente minimiza erros de classificação que poderiam resultar em multas ou autuações fiscais.
- Eficiência Operacional: A classificação automática por tipo de documento e centro de custo, juntamente com a organização em lote, elimina a necessidade de manipulação manual de arquivos, liberando a equipe fiscal para tarefas de maior valor estratégico.
- Inteligência de Negócio: O sistema não apenas classifica, mas também gera alertas setoriais e implicações fiscais específicas. Por exemplo, o agente pode alertar sobre a necessidade de conferir a MVA em operações com Substituição Tributária no Setor Automotivo, ou indicar que um documento de entrada é um insumo para cálculo de custo de produção na Indústria.
- Adaptabilidade a Desafios: A arquitetura de agentes e o CustomizacaoAgent permitem que o sistema se adapte rapidamente a mudanças legais e trate setores não-padrão (como Terceiro Setor ou Órgãos Públicos), um desafio crítico na área fiscal.

4. Detalhamento do que foi Desenvolvido

O desenvolvimento se concentrou na criação de uma arquitetura modular e orientada a agentes, conforme detalhado abaixo:

4.1. Arquitetura de Agentes

Componente	Função Principal	Detalhes da Operação
Agente Orquestrador (orchestrator_agent.py)	Gerencia o fluxo de processamento.	1. Recebe o documento. 2. Delega a extração de dados (XML/PDF). 3. Infere o Ramo de Atividade (via CNAE). 4.

		Chama o CFOPClassifierAgent para classificação base. 5. Delega a análise setorial ao Agente Especializado correspondente.
Agente Classificador Base (cfop_classifier_agent.py)	Realiza a classificação fundamental.	Classifica o documento por Tipo (Compra, Venda, Serviço) e sugere o Centro de Custo com base no CFOP e nas configurações gerais.
Agentes Especializados (Ex: agronegocio_agent.py)	Aplica a inteligência fiscal customizada.	Gera Implicações Fiscais, Recomendações de Arquivamento e Alertas Setoriais específicos para o ramo detectado, utilizando as regras definidas em ramos_atividade.json.
Agente de Customização (customizacao_agent.py)	Lida com exceções e adaptações.	Analisa documentos para identificar setores de tratamento especial (Órgãos Públicos, Terceiro Setor) e gera alertas sobre a necessidade de adaptação a mudanças legais recentes.

4.2. Fluxo de Operação

A solução é operada através de um Dashboard Interativo (Streamlit) que oferece dois modos de uso:

- Análise Individual: O usuário faz o upload de um único arquivo (XML ou PDF). O sistema processa o documento e exibe imediatamente no dashboard:
 1. O Ramo de Atividade detectado.
 2. O Centro de Custo sugerido.
 3. Uma lista detalhada de Implicações Fiscais, Alertas Setoriais e Recomendações de Arquivamento geradas pelos Agentes Especializados.

- Processamento em Lote: O usuário coloca múltiplos arquivos na pasta de entrada (data/notas/) e ação o processamento. O sistema:
 1. Processa cada arquivo, aplicando a lógica de agentes.
 2. Copia o arquivo classificado para uma estrutura de pastas organizada por Ramo de Atividade e Data (output/Ramo/Ano-Mês/).
 3. Exibe um Sumário da Classificação no dashboard, indicando o total de sucessos e falhas, e confirmando a conclusão da organização.

4.3. Manutenção e Expansão

A customização e a adaptabilidade são facilitadas pela manutenção de arquivos de configuração JSON:

- ramos_atividade.json: Permite adicionar novos ramos de atividade ou atualizar as regras (impostos, particularidades) dos ramos existentes sem alterar o código principal.
- cnae_ramo_map.json: Facilita a atualização do mapeamento CNAE-Ramo, garantindo que a detecção inicial do setor esteja sempre alinhada com as necessidades do negócio.
- crawler.py: Garante que a base de dados de CFOPs, essencial para a classificação base, esteja sempre atualizada com a legislação do CONFAZ.

5. Detalhamento do Componente: orchestrator_agent.py

O arquivo orchestrator_agent.py é o núcleo de controle da arquitetura de Agentes Especializados. Sua função principal é coordenar o fluxo de trabalho, desde a recepção do documento até a consolidação da análise fiscal final, garantindo que a inteligência setorial customizada seja aplicada corretamente.

5.1. Funções e Responsabilidades

O OrchestratorAgent é responsável por:

1. Inicialização e Configuração: Carregar todos os agentes especializados e as configurações iniciais (como a base de dados de CFOP e o mapeamento CNAE-Ramo).
2. Orquestração do Fluxo: Gerenciar a sequência de etapas de processamento: Extração -> Detecção de Ramo -> Classificação Base -> Análise Setorial Customizada.
3. Detecção de Ramo de Atividade: Inferir o ramo de atividade da empresa emitente ou destinatária, utilizando o CNAE como fonte primária e o CFOP como fallback.

4. Processamento de Documento Único: Coordenar a análise de um único arquivo (usado pelo Dashboard para análise individual).
5. Processamento em Lote: Gerenciar a análise de múltiplos arquivos, organizando-os na estrutura de pastas de saída.

5.2. Integração com Agentes Especializados (O Coração da Customização)

O ponto mais importante do orchestrator_agent.py é a forma como ele implementa a customização setorial. Após a classificação base, o orquestrador executa a seguinte lógica:

1. Seleção do Agente: O orquestrador usa o ramo detectado (ramo_detectado) para buscar o agente correspondente no dicionário self.agentes_especializados.
2. Delegação da Análise: Se o agente for encontrado, o orquestrador chama o método analisar_documento do agente especializado, passando o CFOP e os dados extraídos.
3. Mesclagem de Resultados: Os resultados da análise setorial (alertas, implicações, recomendações) são mesclados ao resultado da classificação base, garantindo que a resposta final contenha a inteligência customizada.

Além disso, o orquestrador sempre chama o CustomizacaoAgent para adicionar uma camada de análise sobre setores de tratamento especial (órgãos públicos, terceiro setor) e alertas sobre mudanças legais, garantindo que os desafios de adaptação sejam endereçados em todas as operações.

5.3. Como a Solução é Operada (Visão do Orquestrador)

O OrchestratorAgent é instanciado uma única vez pelo Dashboard (dashboard.py). A operação é simples:

- Análise Individual: O Dashboard chama agent.processar_documento(caminho_do_arquivo).
- Processamento em Lote: O Dashboard chama agent.processar_lote_notas().

O orquestrador cuida de toda a complexidade interna, garantindo que a extração, classificação e customização setorial ocorram de forma sequencial e coesa, entregando um resultado final unificado e rico em informações fiscais.

Disponível em: https://github.com/annalug/grupo_i2a2