

INTELIGÊNCIA ARTIFICIAL A SERVIÇO DO FISCO



Ricardo Costa Pinto

- Delegado Regional de Fiscalização de Luziânia
- Auditor-fiscal da Receita Estadual de Goiás
- MBA em Data Science e Analytics pela USP/ESALQ, Especialista em Sistemas Distribuídos, graduado em Ciência da Computação pela Universidade Federal da Bahia – UFBA



Outubro de 2023.

CONTEÚDO DA **APRESENTAÇÃO**

1 - Inteligência Artificial e Administração Tributária

2 - O Modelo de IA para identificação de empresas fantasmas

3 - Aplicação e Resultados da Rede Neural Artificial

4 - Considerações Gerais da Rede Neural Artificial

5 - Considerações sobre o ChatGPT

6 - Desafios das Administrações Tributárias

7 – Reflexão sobre a Inteligência Artificial

The background features a dark blue space filled with numerous small, glowing particles. Several prominent, wavy, translucent blue lines flow across the frame, creating a sense of motion and depth. These lines are composed of fine, parallel paths of light, giving them a shimmering, ethereal appearance. In the upper left corner, a single, larger, bright white circle is visible, resembling a distant star or planet.

Inteligência Artificial e Administração Tributária

Inteligência Artificial

- ✓ A Inteligência Artificial é um campo de estudo multidisciplinar que abrange varias áreas do conhecimento, visando desenvolver sistemas que possam executar tarefas complexas de forma eficiente e autônoma.
- ✓ O que a Inteligência Artificial é capaz de fazer? Reconhecimento de voz e processamento de linguagem natural (PLN); Aprendizado de máquina e análise de dados; Jogos e entretenimento; Robótica; Diagnóstico médico e tratamento personalizado; Automação de processos; Pesquisa científica e inovação, entre outras.
- ✓ ***“Não é IA Só Porque Alguém Diz que É”.***
- ✓ **Cuidado !** Atualmente existe uma cultura de rotular algoritmos inteligentes como IA, os meios de comunicação frequentemente cometem esse equívoco.



Inteligência Artificial

- ✓ Uma dica para tentar identificar se o sistema realmente utiliza Inteligência Artificial, é questionar quais dados foram usados para treinamento e como ele aprende com os dados.
- ✓ O ChatGPT popularizou o uso de inteligência artificial generativa, como recurso para busca e produção de conteúdos em texto através de um chatbot de respostas imediatas.
- ✓ IA generativa é um tipo de sistema capaz de gerar texto, imagens ou outras mídias em resposta a solicitações em linguagem comum.
- ✓ O conjunto de dados de treinamento da Versão 3 do ChatGPT supera 03 terabytes de texto (conhecimento geral, médicos, jurídicos, exemplos de linguagem de programação, entre outros)



Administração Tributária

- ✓ A Administração Tributária é responsável por acompanhar, controlar e fiscalizar o cumprimento das normas que tratam das receitas tributárias, e, o principal desafio, diz respeito ao **combate da sonegação fiscal**.
- ✓ Instituição do Sistema Público de Escrituração Digital – SPED, a forma tradicional de sonegação, venda sem documento fiscal, foi dificultada em razão da obrigatoriedade de emissão de nota fiscal eletrônica em toda cadeia produtiva (indústria, atacado, varejo e serviço).
- ✓ A fiscalização tributária tem observado nos últimos anos um aumento significativo de empresas que estão sendo constituídas somente no papel (empresas fantasmas).



Administração Tributária

- ✓ Necessidade de aprimoramento das técnicas de fiscalização tributária, utilizando ferramentas e tecnologias capazes de identificar padrões de empresas que são constituídas para praticar o crime de sonegação.
- ✓ As Administrações Tributárias via de regra utilizam malhas fiscais e cruzamento de dados para tentarem identificar possíveis fraudes. Esse tipo de metodologia tradicional não emprega técnicas de aprendizado de máquina (machine learning).
- ✓ O presente trabalho possui como objetivo aplicar predição baseada em técnicas de inteligência artificial capaz de identificar padrões típicos de contribuintes desaparecidos.



The background features a dark blue field with glowing, wavy lines of light blue and white particles, creating a sense of motion and digital data. A single, larger white circle is visible in the upper left corner.

O Modelo de Inteligência Artificial para identificação de empresas fantasmas

O modelo de Inteligência Artificial para identificação de empresas fantasmas

- ✓ A técnica de Inteligência Artificial utilizada foi uma Rede Neural Artificial (aprendizado supervisionado), que através dos neurônios artificiais, possui capacidade de predição, assim como o cérebro humano, gerando um resultado mais assertivo para a fiscalização.
- ✓ O modelo de Rede Neural Artificial **classifica** os contribuintes, a partir do conjunto de **25 (vinte e cinco) variáveis explicativas**, como suspensas (desaparecidas) ou ativas. Essas variáveis demonstram características relevantes de cada empresa, tais como: **arrecadação, notas fiscais, venda de cartão de crédito/débito, atividade econômica, localização, porte, área do estabelecimento, fornecedores, sócios, entre outras.**
- ✓ Foram utilizados dados de **27.709** empresas do cadastro de contribuintes do Estado de Goiás, pessoas jurídicas de diversos ramos de atividades econômicas localizadas na região do entorno sul do Distrito Federal.



O modelo de Inteligência Artificial para identificação de empresas fantasmas

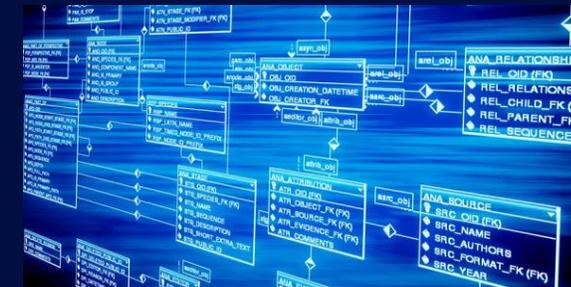
- ✓ Em relação à plataforma tecnológica, foram desenvolvidos scripts específicos na Linguagem R para tratamento dos dados e construção da Rede Neural Artificial.
- ✓ Após o tratamento dos dados, utilizando técnicas de Data Wrangling para estruturação, padronização, organização, limpeza, normalização e validação dos dados, a **rede neural** foi submetida a um processo de **aprendizagem**.
- ✓ O aprendizado (treinamento) da rede neural foi realizado com dados de **18.769** contribuintes, aproximadamente 70% do conjunto total selecionado.
- ✓ Dessa forma é possível determinar com mais assertividade as pessoas jurídicas a serem fiscalizadas in loco, para averiguação/confirmação da fraude praticada.



O modelo de Inteligência Artificial para identificação de empresas fantasmas

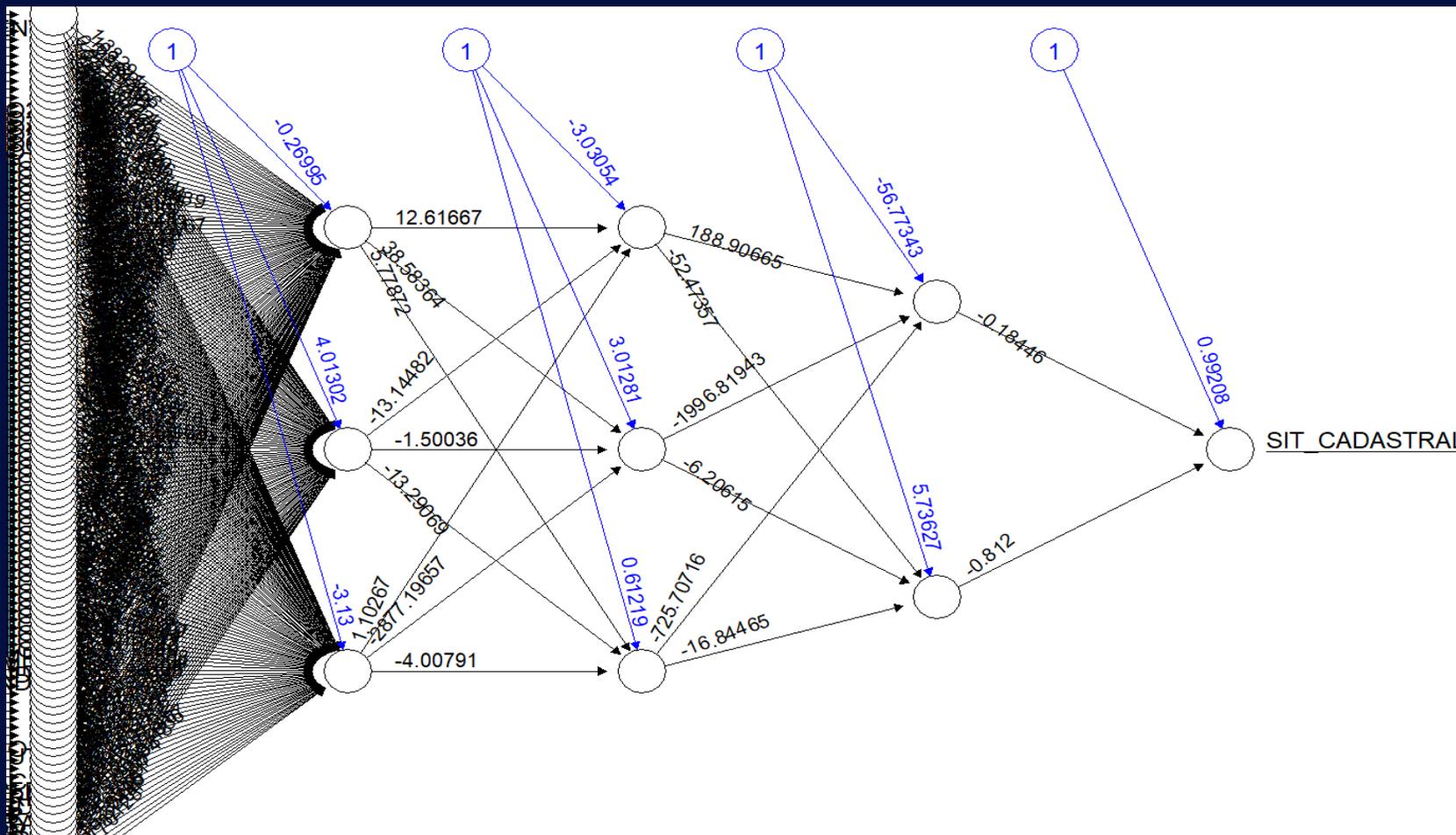
✓ Bases de dados (Dataset's) do modelo.

Environment	History	Connections	Tutorial
Import Dataset 634 MiB			
R			
Data			
ds_arrecadacao	10843 obs. of 5 variables		
ds_contribuinte	27709 obs. of 21 variables		
ds_fat_cartao_credito	413 obs. of 5 variables		
ds_fat_cartao_debito	117 obs. of 5 variables		
ds_info_socio	28019 obs. of 5 variables		
ds_matrix_teste	Large matrix (868752 elements, 7.5 MB)		
ds_matrix_treinamento	Large matrix (2027052 elements, 17.4 MB)		
ds_nfe_compra	12509 obs. of 15 variables		
ds_preprocessamento	27369 obs. of 27 variables		
ds_socio	27007 obs. of 3 variables		
ds_socio_FORA_GO	5271 obs. of 2 variables		
ds_socio_GO	22616 obs. of 2 variables		
ds_split	Large initial_split (4 elements, 4.7 MB)		
ds_teste	8044 obs. of 27 variables		
ds_treinamento	18769 obs. of 27 variables		
normalizacao	List of 21		



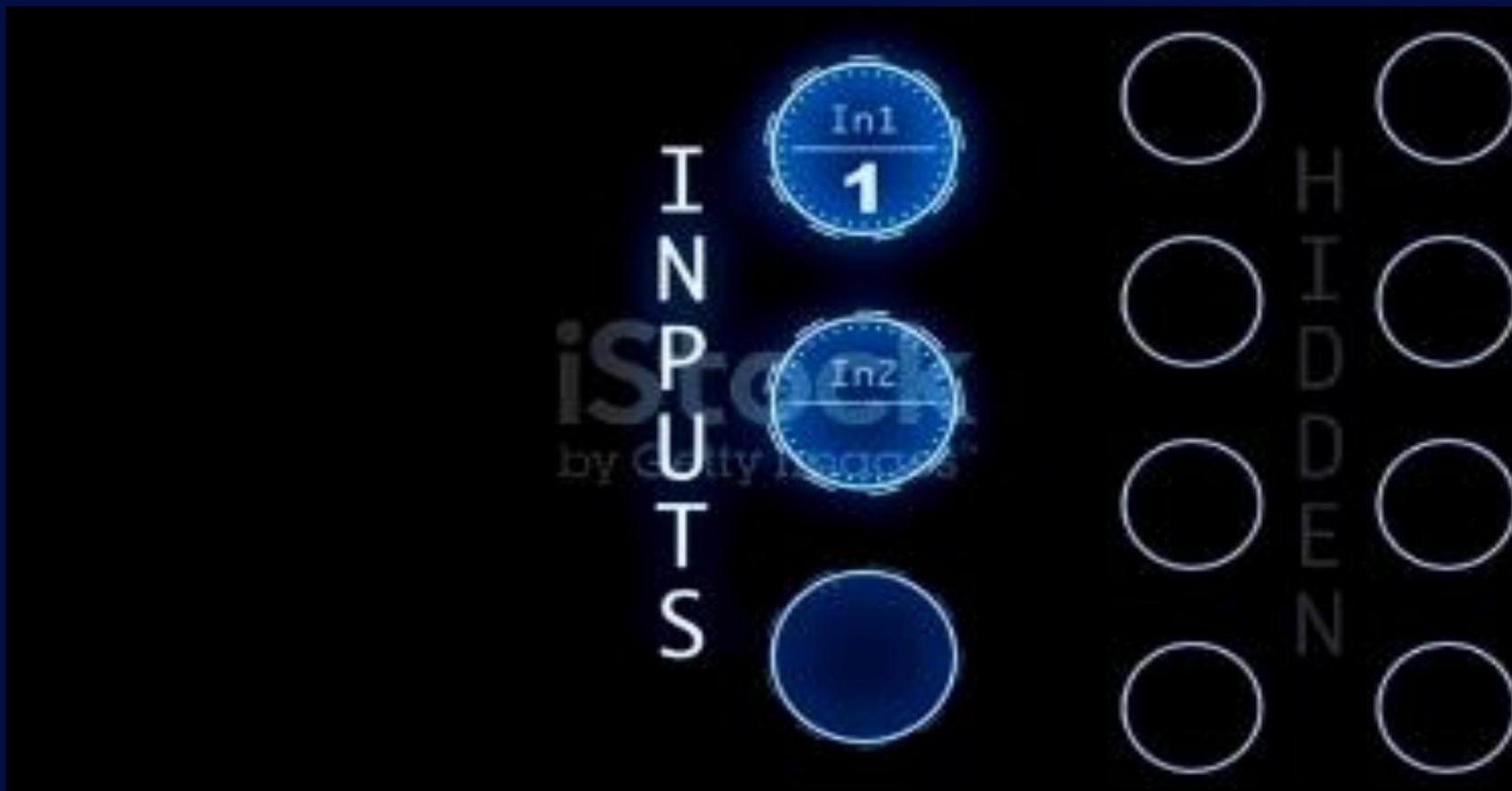
O modelo de Inteligência Artificial para identificação de empresas fantasmas

✓ Visualização Gráfica da RNA – Rede Neural Artificial.



O modelo de Inteligência Artificial para identificação de empresas fantasmas

- ✓ Animação simulando o funcionamento de uma Rede Neural Artificial.

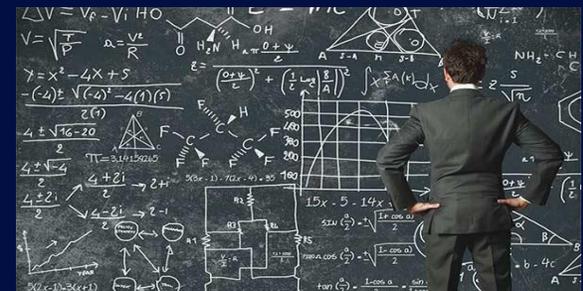


The background features a dark blue space filled with numerous small, glowing particles. Overlaid on this are several large, flowing, wavy lines composed of a grid of fine dots, creating a sense of motion and depth. A single, larger, bright blue circle is positioned in the upper left corner.

Aplicação e Resultados da Rede Neural Artificial

Aplicação e Resultados da Rede Neural Artificial

- ✓ Inicialmente a Rede Neural Artificial avaliou aproximadamente 8 (oito) mil inscrições estaduais, das quais 965 estão **ativas** no cadastro estadual, porém classificadas como **desaparecidas** pela IA.
- ✓ Servidores da Delegacia Regional de Fiscalização de Luziânia foram alocados para realizar vistorias/diligências nos 965 contribuintes selecionados pela rede neural. O objetivo é verificar se as empresas realmente existem de fato.
- ✓ Até o presente momento, 717 contribuintes foram fiscalizados, com a efetivação de 680 suspensões/inabilitações, uma assertividade **surpreendente de 94,8%**. Ou seja, somente 37 empresas estavam realmente funcionando de fato, todas as demais estavam desaparecidas conforme classificação da IA.



The background features a dark blue space filled with a grid of small, glowing blue dots. These dots form several large, flowing, wavy patterns that sweep across the frame from the bottom left towards the top right. The overall effect is that of a dynamic, digital or data-driven environment. In the top left corner, there is a single, larger, bright blue circle.

Considerações Gerais da Rede Neural Artificial

Considerações Gerais da Rede Neural Artificial

- ✓ A rede neural artificial implementada demonstrou que essa tecnologia auxilia a fiscalização tributária na identificação de empresas que não existem de fato, dessa forma coibindo a sonegação fiscal.
- ✓ É perceptível a contribuição para o desenvolvimento de trabalhos futuros com a aplicação do mesmo método, bem como a necessidade de se trabalhar este tipo de modelo com um **número maior de variáveis** explicativas relevantes.
- ✓ O ganho deste trabalho é mais vasto que demonstrado, desse modo é factível ampliar o estudo científico, baseado em técnicas de Inteligência Artificial, voltado para a fiscalização tributária.
- ✓ É notória a carência da aplicação de ferramentas desse tipo nas Administrações Tributárias.



The background is a dark blue space filled with abstract digital elements. It features several wavy, ribbon-like structures composed of fine, glowing blue lines that create a sense of depth and movement. Scattered throughout the scene are numerous small, bright blue particles, some appearing as simple dots and others as slightly larger, more complex shapes. In the upper left corner, there is a single, larger, bright white circle. The overall aesthetic is futuristic and high-tech.

Considerações sobre o ChatGPT

Considerações sobre o ChatGPT

- ✓ O ChatGPT não conseguiu solucionar problemas relacionados à sonegação fiscal, empresas fantasmas, empresas noteiras quando questionado de *forma genérica*.

INTERAÇÃO COM CHATGPT:

EU: *“Como criar uma IA para identificar sonegação fiscal do ICMS”*

CHATGPT: *“Criar uma IA (Inteligência Artificial) para identificar sonegação fiscal do ICMS (Imposto sobre Circulação de Mercadorias e Serviços) é um projeto complexo que envolve várias etapas. Abaixo, apresento um guia geral para ajudar você a iniciar esse processo: Entenda as regras fiscais do ICMS, Dados disponíveis, Pré-processamento de dados, Escolha da técnica de IA,”*

EU: *“Você respondeu superficialmente a pergunta.”*

CHATGPT: *“Peço desculpas se a resposta anterior não foi tão detalhada como você esperava. Vou fornecer uma explicação mais aprofundada sobre como criar uma IA para identificar empresas noteiras: Compreender o problema, Coleta e preparação de dados, Engenharia de recursos, Escolha do modelo, Conjunto de treinamento, Treinamento do modelo, Avaliação do modelo, Ajustes e otimizações, Implementação e integração e Monitoramento contínuo.”*



Considerações sobre o ChatGPT

INTERAÇÃO COM CHATGPT:

EU: “Pelo visto vocês IA precisam aprender muito em relação a sonegação de impostos, identificação de empresas fantasmas e noteiras. O conhecimento de vocês é muito superficial nesse assunto, rasteiro.”

CHATGPT: “Peço desculpas se minhas respostas não foram tão detalhadas ou abrangentes quanto você esperava. Como uma IA baseada em texto, meu conhecimento é limitado ao que foi fornecido nos dados de treinamento até minha data de corte em setembro de 2021. Embora eu tenha uma ampla base de conhecimento em várias áreas, incluindo sonegação de impostos e práticas fraudulentas, é importante reconhecer que nenhum sistema de IA é perfeito e que ainda há muito a aprender e melhorar.

A identificação de empresas fantasmas, noteiras e práticas de sonegação fiscal é um assunto complexo e multifacetado, envolvendo questões legais, contábeis, fiscais e de inteligência investigativa. É importante destacar que a luta contra essas práticas requer esforços conjuntos de especialistas humanos, agências governamentais, órgãos de fiscalização e avanços contínuos em tecnologia, incluindo inteligência artificial. “



Considerações sobre o ChatGPT

INTERAÇÃO COM CHATGPT:

EU: “Quais variáveis relevantes devo selecionar para criar uma rede neural artificial que identifica empresas fantasmas ou empresas noteiras ?.”

CHATGPT: “A criação de uma rede neural artificial (RNA) para identificar empresas fantasmas ou noteiras é um **desafio complexo** e requer a seleção cuidadosa de variáveis relevantes. As variáveis que você deve escolher dependem das informações disponíveis nos seus dados e da natureza específica do problema. No entanto, aqui estão algumas variáveis que **geralmente** são relevantes para identificar empresas fantasmas ou noteiras:

1. Atividade Comercial: Setor de Atuação, Histórico de Transações
2. Dados Financeiros: Receita e Lucro, Fluxo de Caixa, Dívidas e Passivos
3. Registro Legal: Data de Fundação, Registros de Propriedade
4. Compliance Fiscal: Regularidade Tributária, Histórico de Notas Fiscais
5. Características Legais: Endereço Comercial, Registros Legais
6. Redes e Conexões: Redes de Empresas, Registros de Contas Bancárias
7. Comportamento de Pagamento: Histórico de Pagamentos
8. Anomalias Estatísticas: Anomalias em Dados Financeiros
9. Sentimento Online: Reputação Online
10. Dados Comparativos: Benchmarking

É importante notar que não existe uma lista definitiva de variáveis, e a escolha final dependerá da disponibilidade e qualidade dos dados, bem como da **expertise em análise de fraude tributária e fiscal**. Além disso, você pode considerar o uso de técnicas de aprendizado de máquina, como árvores de decisão, redes neurais, ou algoritmos de detecção de anomalias, para criar um modelo que leve em consideração essas variáveis e identifique empresas fantasmas ou noteiras com base nos padrões observados nos dados.”





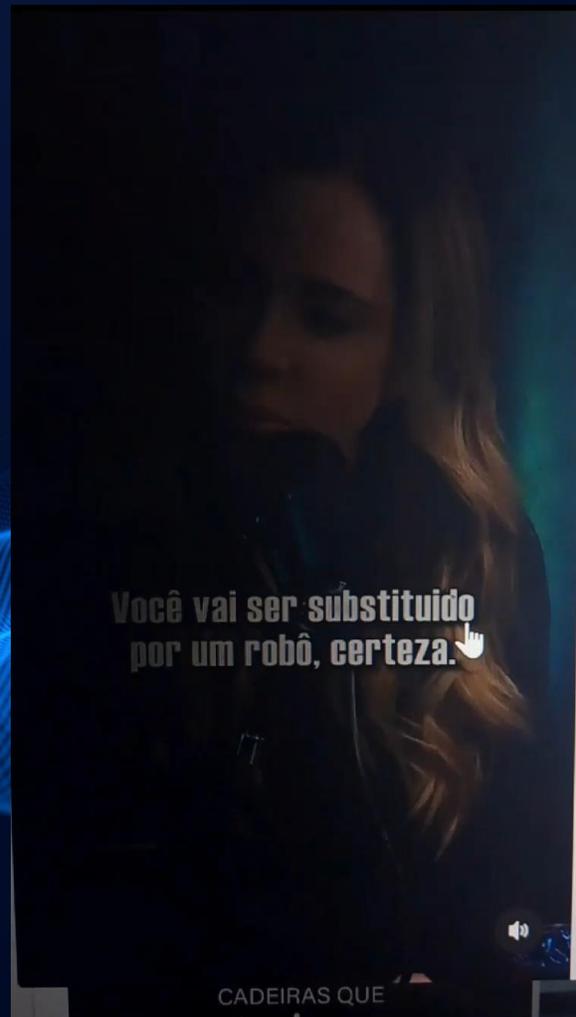
Desafios das Administrações Tributárias

Desafios das Administrações Tributárias

- ✓ Aprimorar a eficiência da Administração Tributária através do uso de técnicas de Inteligência Artificial nas diversas áreas.
- ✓ Criar base de conhecimento para implementação de novos modelos de Inteligência Artificial - IA.
- ✓ Criar novos modelos de IA para auxiliar as Administrações Tributárias, dentre eles: identificação de operações fiscais e contábeis fraudulentas; projeção de receitas tributárias; recuperação de crédito; etc.
- ✓ Investir em pesquisa e capacitação nas áreas de ciência de dados e aprendizado de máquina (machine learning).



Reflexão sobre Inteligência Artificial



"Existe apenas um canto do universo que você pode ter certeza de aperfeiçoar, que é você mesmo." (Aldous Huxley)

OBRIGADO!

Ricardo Costa Pinto

 (61) 99986-5764

 ricardo.rcosta@goias.gov.br