

# Victor Taendy Sousa Emerenciano

São Paulo - SP  
(11) 9 5377-5093  
victor338.flow@gmail.com  
<https://www.linkedin.com/in/victor-taendy/>

Cientista de Dados com sólida experiência em desenvolvimento de soluções baseadas em aprendizado de máquina, processamento de imagens e IA generativa. Com foco na entrega de resultados mensuráveis e impacto direto nos processos de negócio.

---

## Habilidades

Python, Pandas, Scikit-Learn, TensorFlow, Pytorch | MS SQL Server, Hadoop Hive, AWS Athena, SageMaker Studio | OpenCV, NLTK, SpaCy, XGBoost, LightGBM | Git, GitHub | Análise de dados

---

## Experiência

### 2023 – 2025 Cientista de Dados Pleno – Itaú Unibanco

Monitoramento de Ambientes com Imagens

- Implementação de modelos de detecção de entidades (YOLO) para identificar presença humana em ambientes específicos, utilizando soluções da Ultralytics e OpenCV.
- Desenvolvimento de modelo de classificação com PyTorch para atuar como filtro preliminar, reduzindo a carga computacional dos modelos de detecção.

Priorização de Operações para Quality Assurance

- Desenvolvimento de modelo de classificação utilizando Naive Bayes para melhorar a seleção de operações críticas para análise, superando a abordagem de amostragem aleatória em 0,5 ponto percentual na métrica Average Precision.
- Criação de visualizações para comunicar efetivamente os ganhos do modelo em termos de recall e precisão para stakeholders não técnicos.

Extração de informações não padronizadas com IA generativa

- Desenvolvi uma solução baseada em IA generativa para extrair informações críticas de contratos societários, caracterizados por linguagem natural e estrutura não padronizada.
- Superando as limitações dos métodos tradicionais de detecção de entidades, resultando em maior agilidade na atualização de cenários empresariais e eficiência operacional.

### 2021 – 2022 Cientista de Dados Junior – Itaú Unibanco

Classificação de alto desbalanceio para operações suspeitas

- Desenvolvi um modelo de ranqueamento (XGBoost “learn-to-rank”) para priorizar operações suspeitas em um cenário com alto desbalanceamento de classes, elevando a métrica de 5% com técnicas anteriores baseadas em Random Forest para 11% com o modelo atual.
  - Essa melhoria permitiu uma triagem mais eficiente das transações, mantendo o custo com fraudes igual mesmo com um aumento substancial de tentativas de ataque.
- 

## Formação

2014 – 2020: Bacharelado em Sistemas de Informação  
Escola de Artes Ciências e Humanidades da Universidade de São Paulo