

Jeová Anderson

jeova.herminio@gmail.com | +55 83 98803-8348

- Projetos & Portifolio | <https://geovatatsuga.github.io/portifolio/>
- LinkedIn | <https://www.linkedin.com/in/jeova-anderson/>
- Medium | <https://medium.com/@jeova.anderson>
- Github | <https://github.com/geovatatsuga>

OBJETIVO PROFISSIONAL

Analista de Dados Júnior | Cientista de Dados Júnior | Engenharia de Dados Júnior | BI / Machine Learning

RESUMO PROFISSIONAL

Estudante de Ciência de Dados (UEPB) com experiência prática em Machine Learning, Engenharia de Dados, análise de dados e construção de dashboards. Atuo com Python, SQL, PySpark, Azure Databricks, Delta Lake, MLflow, Power BI, Looker Studio, Dash e Plotly para criação de pipelines, modelos preditivos, análises de negócio e visualizações executivas. Experiência com IA Generativa, RAG e chatbots, além de inglês técnico sólido para leitura de documentação, escrita de código e comunicação em ambientes internacionais.

FORMAÇÃO ACADÊMICA

Tecnólogo Ciência de Dados - Universidade Federal da Paraíba - Conclusão 06/2026

Bolsa integral na UEPB por meio de processo seletivo e transição para continuidade dos estudos com foco em excelência acadêmica.

EXPERIÊNCIAS PROFISSIONAIS

Vsoft Tecnologia - Estagiário em Governança Corporativa - 07/2025 - 12/2025

- Desenvolvi e mantive sistema de RAG para consulta de documentação interna, reduzindo em ~40% o tempo de busca manual por informações.
- Participei da criação do Chatbot Vivi, automatizando mais de 200 interações mensais de usuários internos.
- Automatizei o envio de e-mails de gestão de contratos via scripts Python, eliminando cerca de 3h semanais de trabalho manual.
- Construí dashboard integrada consolidando KPIs de 3 áreas, adotada pela liderança como ferramenta de monitoramento mensal.

Arquivo dos Governadores - Estagiário - 01/2024 - 12/2024

- Organizei e digitalizei mais de 1.200 documentos históricos, viabilizando acesso digital ao acervo.
- Criei tabelas dinâmicas no Excel que reduziram em ~50% o tempo de sistematização e controle documental.
- Reestruturei o fluxo de contagem e classificação de acervos, aumentando a produtividade do setor em aproximadamente 30%.

Marcos Inácio Advogados - Auxiliar Administrativo - 04/2021 - 08/2021

- Controlei entrada e saída de materiais e emissão de recibos, mantendo acurácia de 100% nos registros durante o período.
- Implementei planilha de controle de recursos que reduziu em ~25% o tempo gasto na gestão administrativa do escritório.

PROJETOS

Pipeline de Detecção de Fraudes com Azure Databricks + Delta Lake - 05/2026

[Medium](#) | [GitHub](#)

- Construção de pipeline lakehouse no Azure Databricks para detecção de fraudes financeiras no dataset PaySim, processando mais de 6,36 milhões de transações.
- Organização dos dados em camadas Bronze, Silver e Gold com PySpark, Delta Lake e ADLS Gen2.
- Engenharia de features comportamentais usando erros de saldo, tipo de transação, saldo zerado e transformação logarítmica do valor.
- Treinamento e comparação de Logistic Regression, Decision Tree, Random Forest e GBT com rastreamento no MLflow.
- Random Forest com features definido como modelo recomendado, alcançando 99,45% de recall e ganho financeiro simulado de R\$ 73,21 milhões.

Dashboard e Previsão de Churn com Random Forest - 04/2025

[Medium](#) | [GitHub](#)

- Criação de solução integrada com coleta de dados, modelo preditivo e dashboards para previsão e acompanhamento de churn.
- Extração e pré-processamento de dados com SQL Server, Python, Pandas e pyodbc.
- Treinamento de modelo Random Forest em 1.926 amostras, alcançando 81% de acurácia geral.
- Identificação de variáveis relevantes para churn, com destaque para tipo de contrato, método de pagamento e receita total.
- Construção de dashboards em Dash e Plotly para listar clientes com maior risco e apoiar estratégias preventivas de retenção.

Segmentação de Clientes com RFM + K-Means - 04/2025

[Medium](#) | [GitHub](#)

- Criação de segmentação comportamental baseada em Recência, Frequência e Monetariedade para classificar clientes por valor e engajamento.
- Classificação dos clientes em 10 subgrupos comportamentais, como adormecidos, em risco, parceiros de longa data e grupos de maior valor.
- Aplicação de transformações estatísticas como logarítmica, raiz quadrada e Box-Cox para reduzir assimetria dos dados.
- Definição de 4 clusters com K-Means após análise do número ideal de agrupamentos.
- Desenvolvimento de dashboard interativo em Dash e Plotly para apoiar ações de marketing, vendas e retenção.

Detecção de Imagens Sintéticas com ViT + EBP - 12/2025

[Medium](#) | [GitHub](#)

- Desenvolvimento de abordagem híbrida combinando Vision Transformer com Edge-Based Processing para classificação de imagens reais vs sintéticas.
- Fine-tuning de ViT ImgNet21k com PyTorch e HuggingFace Transformers, usando otimizações para GPU.
- Implementação de módulo EBP para análise estrutural baseada em bordas e altas frequências.
- Alcance de 99,43% de acurácia e F1-Score de 0,9943 com a combinação ViT Fine-tuned + EBP.
- Melhoria do modelo ViT Fine-tuned de 98,89% para 99,43%, reduzindo erro residual em cenário de alta performance.

Criando e Aplicando Ranqueamento por ELO Rating no Jiu-Jitsu - 06/2025

[Medium](#) | [GitHub](#)

- Construção de sistema de ranking baseado em ELO para medir desempenho competitivo de atletas de Jiu-Jitsu ao longo do tempo.
- Coleta, limpeza e estruturação de mais de 56 mil registros de lutas com Python, BeautifulSoup, ThreadPoolExecutor e Pandas.
- Identificação de picos históricos de performance, com Nicholas Meregali atingindo 2.485 pontos de Peak ELO.
- Aplicação de K-Means para separar atletas em 3 perfis competitivos, incluindo grupo de elite com 154 atletas.
- Criação de visualizações estatísticas para análise de evolução de ELO, taxa de vitória, volume de lutas e padrões competitivos.

CURSOS COMPLEMENTARES

Data Literacy - IBM SkillsBuild

Credencial: <https://www.credly.com/go/5uEt7GGF> | Emitido em: 28 ABR 2026

2026 Aspire Leaders Program - Aspire Institute

Certificado de conclusão de 40 horas | Abril 2026

Artificial Intelligence Fundamentals - IBM

Credencial: 9768a0cb-ca98-4caf-8f03-455b20edbe07

Fundamentos de Data Science e Inteligência Artificial - Data Science Academy

Credencial: 672502adfbadd182d908fdad

Lei Geral de Proteção de Dados Pessoais - Fundação Bradesco

Credencial: 01A397D8-1221-4875-8403-43EF6A42C948

PRINCIPAIS FERRAMENTAS

Linguagens: Python, SQL.

Dados e Engenharia: PySpark, Delta Lake, Azure Databricks, ADLS Gen2, ETL, APIs, Web Scraping, Pandas, NumPy.

Machine Learning: Scikit-learn, MLflow, Random Forest, GBT, Decision Tree, Logistic Regression, K-Means, modelagem preditiva e clusterização.

Deep Learning e IA: PyTorch, HuggingFace Transformers, Vision Transformer, RAG, LLMs, Chatbots.

Visualização e BI: Power BI, Looker Studio, Dash, Plotly, Streamlit, Matplotlib, Seaborn, Excel.

Cloud e Ferramentas: Azure, Git, GitHub, SQL Server, Google Colab.

IDIOMAS

Português: Nativo.

Inglês: Técnico avançado - leitura de documentação, escrita de código e comunicação em ambientes internacionais. Conversacional intermediário.

Espanhol: Básico (leitura e compreensão).

DIFERENCIAIS

- Altamente adaptável e domínio de diferentes ferramentas, linguagens e stacks de dados.
- Experiência prática com Inteligências Artificiais Generativas, RAG, chatbots e soluções inovadoras baseadas em dados.
- Capacidade de compreender o contexto de negócio e transformar análises técnicas em impacto real.
- Habilidade de comunicar insights para públicos técnicos e não técnicos.
- Comprometimento com aprendizado contínuo e crescimento profissional constante.

+ SOBRE MIM

Movido por curiosidade, tecnologia e vontade de resolver problemas reais, busco evoluir constantemente em projetos que unam dados, engenharia, inteligência artificial e impacto de negócio. Tenho satisfação em transformar dados em soluções práticas, seja por meio de pipelines, modelos preditivos, dashboards ou automações. Acredito que o crescimento em Ciência de Dados vem da combinação entre prática, colaboração, documentação e aprendizado contínuo.