

SUAMI EVELIN ROCHA DE MEDEIROS

Engenheira de Dados Júnior | Analista de Dados | Business Intelligence

Araguaína, Tocantins, Brasil +55 63 99275-4343

suamirochat@gmail.com

linkedin.com/in/surocham

github.com/surocham

surocham.github.io/srm-portfolio

RESUMO PROFISSIONAL

Engenheira de Dados com sólida formação em Análise e Desenvolvimento de Sistemas e especialização em Data Analytics. Experiência comprovada na construção de pipelines ETL/ELT escaláveis, processamento distribuído de grandes volumes de dados e desenvolvimento de soluções analíticas orientadas a negócios. Domínio de tecnologias cloud-native, orquestração de workflows complexos e implementação de arquiteturas de dados modernas. Comprometida com a entrega de soluções de alta performance que impulsionam decisões estratégicas baseadas em dados.

COMPETÊNCIAS TÉCNICAS

Engenharia de Dados: Pipeline ETL/ELT, Data Warehousing, Data Lake Architecture, Data Modeling, Data Quality, Data Governance, Batch Processing, Stream Processing, Workflow Orchestration

Linguagens de Programação: Python (Avançado), SQL (Intermediário), Bash Scripting

Frameworks e Bibliotecas: Pandas, NumPy, PySpark, PyTorch, SQLAlchemy, Apache Spark

Big Data e Streaming: Apache Spark, Apache Kafka, Apache Airflow, Hadoop Ecosystem

Cloud Platforms: AWS (S3, Glue, Redshift, Lambda), Azure (Data Factory, Synapse), Google Cloud (BigQuery), Databricks

DevOps & Infraestrutura: Docker, Git/GitHub, CI/CD Concepts, Infrastructure as Code, Container Orchestration

Business Intelligence: Power BI, Data Visualization, Dashboard Development, KPI Tracking, Excel Advanced

Bancos de Dados: PostgreSQL, MySQL, Amazon Redshift, Google BigQuery, Data Warehouse Design

EXPERIÊNCIA PROFISSIONAL

Estagiária em Desenvolvimento Mobile | Branvier

Belém, Pará | Fevereiro 2024 – Outubro 2024

Desenvolveu e implementou pipeline analítico para monitoramento de métricas de aplicação mobile, processando dados de mais de 50.000 usuários ativos mensais

Estruturou modelo dimensional para análise de comportamento de usuários, reduzindo tempo de geração de relatórios em 35%

Automatizou processo de coleta e transformação de dados operacionais, eliminando 15 horas semanais de trabalho manual

Implementou dashboards interativos em Power BI para acompanhamento de KPIs críticos, suportando decisões de produto

Colaborou com equipe multifuncional em ambiente ágil, participando de sprints quinzenais e cerimônias Scrum

Principais Resultados:

Entregou mais de 10 funcionalidades analíticas com impacto direto no produto

Reduziu taxa de erros recorrentes em aproximadamente 20% através de validação automatizada de dados

Suporou 8 releases bem-sucedidas baseadas em insights orientados por dados

PROJETOS EM ENGENHARIA DE DADOS

Pipeline ETL – Análise Epidemiológica de Dengue no Brasil

Desenvolvimento Contínuo | 2025

Arquitetou e implementou pipeline ETL robusto para processamento de dados epidemiológicos de saúde pública, manipulando datasets com mais de 100.000 registros históricos de casos de dengue.

Responsabilidades e Realizações:

Desenvolveu módulos Python para ingestão automatizada de dados públicos de fontes governamentais

Implementou rotinas de data cleaning e normalização, tratando inconsistências e valores ausentes em datasets complexos

Projetou modelo dimensional analítico otimizado para queries de agregação temporal e geográfica

Executou transformações SQL complexas para geração de métricas epidemiológicas e indicadores de tendência

Criou datasets analíticos prontos para consumo por ferramentas de visualização e análise estatística

Stack Tecnológico: Python, Pandas, SQL, NumPy, Data Transformation

Repositório: github.com/surocham/dengue_no_brasil

Pipeline de Machine Learning – Classificação de Dígitos Manuscritos

Projeto Prático | 2025

Construiu pipeline end-to-end de dados para treinamento e deployment de modelo de deep learning, processando dataset MNIST com 70.000 imagens.

Responsabilidades e Realizações:

Desenvolveu pipeline automatizado de preparação de dados com validação de qualidade e transformações

Implementou modelo de rede neural convolucional utilizando PyTorch com accuracy superior a 95%

Otimizou processo de treinamento, reduzindo tempo de execução em 40% através de batch processing eficiente

Automatizou workflow de inferência para classificação em lote de novas amostras

Stack Tecnológico: Python, PyTorch, NumPy, Deep Learning, Computer Vision

Pipeline Orquestrado com Apache Airflow

Projeto Prático | 2025

Desenvolveu arquitetura de orquestração de pipelines de dados utilizando Apache Airflow em ambiente containerizado.

Responsabilidades e Realizações:

- Criou DAGs complexas para orquestração de workflows ETL multi-stage com gerenciamento de dependências
- Implementou containers Docker para isolamento de ambientes e portabilidade de pipelines
- Configurou schedulers para execução automatizada de jobs em intervalos definidos
- Desenvolveu estratégias de retry e error handling para garantir resiliência do pipeline
- Simulou cenários reais de ingestão, transformação e carga de dados analíticos

Stack Tecnológico: Apache Airflow, Docker, Python, Workflow Orchestration, DAG Design

FORMAÇÃO ACADÊMICA

Pós-Graduação em Data Analytics

FIAP | 2025

Especialização avançada em soluções analíticas enterprise, machine learning aplicado, big data analytics e engenharia de dados moderna. Foco em arquiteturas escaláveis e tomada de decisão orientada por dados.

Disciplinas Principais: Big Data Analytics • Machine Learning for Business • Analytical Data Engineering • Advanced Analytics Solutions • Cloud Data Platforms

Tecnólogo em Análise e Desenvolvimento de Sistemas

UNINTER | 2022 – 2025

Formação completa em desenvolvimento de software, arquitetura de sistemas, modelagem e persistência de dados, engenharia de software e programação avançada.

Disciplinas Principais: Data Modeling • Software Engineering • Advanced Programming • Database Systems • System Architecture

CERTIFICAÇÕES e CURSOS COMPLEMENTARES

Fundamentos de Engenharia de Dados – Data Science Academy

Databricks Lakehouse Fundamentals

IBM Data Engineering Fundamentals

Power BI para Business Intelligence e Data Analytics

Ciência de Dados com Python – Santander Open Academy

HABILIDADES COMPORTAMENTAIS

Comunicação Efetiva, Trabalho em Equipe, Resolução de Problemas Complexos, Pensamento Analítico, Aprendizado Contínuo, Adaptabilidade, Atenção aos Detalhes, Gestão de Tempo, Orientação a Resultados, Colaboração

Multifuncional

IDIOMAS

Português: Nativo

Inglês: Intermediário (Leitura técnica avançada)