

Lis Ramos Barreto de Almeida

Engenheira de Dados

Aracaju - SE

+55 (79) 999212447

lis.barreto.py@gmail.com

[LinkedIn](#)

[GitHub](#)

Perfil

Engenheira de Dados com mais de 5 anos de experiência em desenvolvimento de soluções com Python, focada na otimização de pipelines e processamento de grandes volumes de dados. Atuou em projetos no setor financeiro, incluindo geração de leads de Renda Variável, desenvolvimento de feature store para cartões de crédito e orquestração de modelos online. Especialista em DataOps e Spark Performance Tuning, com experiência em migração de pipelines para redução de custos. Atualmente, está aprimorando conhecimentos em MLOps.

Especialidades: Engenharia de Dados, DataOps, MLOps

Experiência

Engenheira de Dados | Itaú Unibanco | 07/2023 - atualmente

Observabilidade de Infraestrutura e Aplicações | Shared Experience BBA

Desenvolvimento de dashboards no Grafana para monitoramento de métricas de infraestrutura e desempenho do Spark. O acompanhamento permite identificar gargalos, otimizar recursos e melhorar a escalabilidade do sistema.

Principais tecnologias: Apache Spark, Grafana, AWS CloudWatch Metrics, AWS EMR

Otimização de Ingestão de Dados do GA4 com Spark | Shared Experience BBA

Atualmente, estou trabalhando com dados do Google Analytics 4 (GA4), que geram cerca de 100 milhões de registros diários, focando na otimização do processo de ingestão desses dados para alimentar modelos de recomendação. Utilizando técnicas de Spark Performance Tuning, reduzi o tempo de processamento mais lento de 1h30min para apenas 14 minutos, melhorando significativamente a eficiência do sistema.

Principais tecnologias: Apache Spark, AWS EMR, AWS Step Functions

Orquestração de Modelos de Machine Learning Online | Prevenção à Fraudes

Desenvolvimento de uma API utilizando FastAPI para gerenciar o tráfego entre diferentes modelos de Machine Learning online. O projeto incluiu testes rigorosos de

performance e integração, garantindo que a API suportasse 600 requisições por segundo com latência inferior a 50 ms por resposta, assegurando alta eficiência e baixa latência no processamento.

Principais tecnologias: AWS ECS, Docker, Python, FastAPI

Feature Store de Cartão de Crédito | Prevenção à Fraudes

Desenvolvimento de uma feature store de habitualidade transacional para cartão de crédito, abrangendo mais de 30 milhões de registros e 150 variáveis. O projeto exigiu uma colaboração intensa com o time de Ciência de Dados para definir a estratégia de janelas temporais e a criação das features, garantindo a otimização dos dados para modelos preditivos.

Principais tecnologias: Apache Spark, AWS Glue, AWS Step Functions, AWS Athena

Engenheira de Dados | Cora | 04/2023 - 06/2023

Criptografia de Dados | Plataforma de Dados

Trabalhei com algoritmos criptográficos em Python para garantir que dados já criptografados fossem acessíveis apenas por pessoas autorizadas. Minha atuação envolveu a implementação de técnicas de criptografia avançada e a criação de mecanismos que assegurassem a proteção e o controle de acesso, garantindo a segurança e privacidade dos dados.

Principais tecnologias: Python, AWS Lambda, AWS S3, AWS KMS.

Engenheira de Dados | XP Inc. | 10/2022 - 03/2023

Processamento de Dados em Near Real Time | Open Finance

Desenvolvimento de pipelines em near real time para processar dados fornecidos pela API de Open Finance. O sistema permitia atualizações quase que instantâneas, monitoramento contínuo de transações e geração de insights, garantindo escalabilidade e eficiência no tratamento de grandes volumes de dados.

Principais tecnologias: Apache Spark, Apache Kafka, Apache Airflow

Eficiência e Otimização em Pipelines de Dados | Conta Digital

Realizei a migração de pipelines de dados do Databricks para o Airflow, reduzindo significativamente os custos. Com a utilização de Spark no Kubernetes, foi possível otimizar o uso de recursos, e, através de técnicas de Spark Performance Tuning, aumentar a eficiência do processamento de dados.

Principais tecnologias: Apache Spark, Apache Airflow, Grafana, Kubernetes

Processamento de Dados em Streaming | Conta Digital

Implementei pipelines de dados em streaming para a Conta Digital, utilizando Apache Kafka e Spark Streaming. O sistema processava dados em tempo real, garantindo alta disponibilidade e escalabilidade, permitindo a tomada de decisões com base em informações atualizadas instantaneamente, melhorando a eficiência operacional e a experiência do usuário.

Principais tecnologias: Apache Spark, Apache Kafka, Apache Airflow

Estagiária de Dados | XP Inc. | 03/2021 - 09/2022

Participei do programa de estágio XData Journey, que ofereceu experiências em diversas áreas de dados da XP através de job rotation, permitindo identificar a área de maior afinidade ao final.

Ingestão, Qualidade e Governança de Dados | Data Engineering

Minhas principais responsabilidades incluíram garantir a qualidade dos dados com verificações rigorosas e sistemas de alerta, desenvolver estratégias eficazes de linhagem de dados. Também desempenhei funções essenciais de engenharia de dados, como armazenamento eficiente, ingestão de dados em tempo real e orquestração de pipelines escaláveis.

Principais tecnologias: Apache Spark, Apache Airflow, Microsoft Azure, Databricks, Git, Python, SQL, Apache Kafka, Jupyter Notebooks.

Churn de Cartão de Crédito | Business Analytics

Participei ativamente do desenvolvimento do caso de churn de cartão de crédito no XP Bank. O processo envolveu a criação da regra de churn, uma análise detalhada das principais variáveis, além do design, refinamento e monitoramento de um dashboard de churn para acompanhar métricas e padrões relevantes. Com isso, foi possível melhorar o acompanhamento e a tomada de decisões relacionadas ao churn de clientes.

Principais tecnologias: Apache Spark, Power BI, Databricks, Python, SQL

Geração de Leads de Renda Variável | Data Science

Atuei em um projeto de ponta a ponta para geração de leads na mesa de Renda Variável da Rico, utilizando técnicas de modelagem estatística e advanced analytics. Desenvolvi filtros avançados para identificação de leads potenciais e criei listas priorizadas por score. Também implementei pipelines para otimizar o fluxo de dados e automatizei o envio das listas para os operadores, aumentando a eficiência e precisão na identificação de oportunidades de investimento.

Principais tecnologias: Apache Spark, Apache Airflow, Kedro, Microsoft Azure, Databricks, Git, Python, SQL, Jupyter Notebooks

Estudante Pesquisadora | Universidade Federal de Sergipe | 2018 - 2021

Aplicação do Método de Runge-Kutta na Solução das Equações de Adirovitch

A implementação da simulação foi feita em Python, utilizando métodos numéricos para resolver as equações, além de técnicas de visualização e análise de dados. Também foram desenvolvidos algoritmos complexos para garantir precisão e eficiência na simulação.

Principais tecnologias: Python, SciPy, Numpy, Matplotlib, Plotly, Streamlit

Diretora em Empresa Jr. | Universidade Federal de Sergipe | 2018 - 2019

Gerenciamento de Projetos | Physics Jr.

Tive experiência em criar e implementar projetos para promover a disseminação científica e inovação tecnológica, monitorar e manter a qualidade de projetos em andamento, gerenciar contratos com parceiros terceirizados e fornecer consultoria para o desenvolvimento de projetos e planos de ação.

Professora de Física | Universidade Federal de Sergipe | 2015 - 2018

Iniciação à Docência

Particpei do Programa Institucional de Bolsa de Iniciação à Docência (PIBID), voltado para estudantes de graduação no início do curso. Fui inserida no ambiente escolar para observar e refletir sobre as práticas pedagógicas do cotidiano de escolas públicas de educação básica, acompanhado por um professor da escola (supervisor) e um docente da UFS.

Educação

FIAP | 2024 - 2025

Especialização em Engenharia de Machine Learning

PUC Minas | 2023 - 2024

Especialização em DevOps & Continuous Software Engineering

Estácio de Sá | 2021 - 2023

Tecnólogo em Banco de Dados

Universidade Federal de Sergipe | 2020 - 2023

Bacharelado em Engenharia de Computação

Obs.: Transferência para Banco de Dados

Universidade Federal de Sergipe | 2014 - 2020

Licenciatura em Física

Obs.: Transferência para Engenharia de Computação

Prêmios e Reconhecimentos

2023 - Reconhecimento "50 to Follow" - Dados & BI | Gama Academy

2021 - Prêmio Colaborador Revelação | XP Inc.

Cursos e Certificados

2024 - Performance Tuning na Prática | Engenharia de Dados Academy

2024 - Arquitetura de Resiliência: Alteração de Demanda | Itaú Unibanco

2024 - SRE - Trained | Itaú Unibanco

2023 - Professional - Data Engineering | Itaú Unibanco

2023 - Associate - Machine Learning Engineering | Itaú Unibanco

2023 - AWS Certified Cloud Practitioner | Amazon Web Services (AWS)

2023 - ChatGPT Prompt Engineering for Developers | DeepLearning.AI

2022 - Learner Tech XP Inc. - Azure | XP Inc.

2022 - Astronomer Certification for Apache Airflow Fundamentals | Astronomer

2022 - Apache Spark Programming | Databricks

2022 - Introduction to Apache Spark Architecture | Databricks

2021 - Data Engineer Path | Dataquest.io

2021 - Data AI: Introduction to Advanced Analytics and AI in Azure | Microsoft

2021 - Data Science XP | Data Bootcamp

2020 - Kanban Foundation (KIKF) | Certiprof

2020 - Open Source Software Development Methods | The Linux Foundation

2020 - Scrum Foundations Professional Certificate (SFPC) | Certiprof

2020 - DL 101 - Curso Geral de Propriedade Intelectual | WIPO

Participação e Organização de conferências

2022 - Organização da Python Nordeste | Comunidade Python Brasil

2021 - Entrevista Mulheres em Dados na XP Inc. | Podcast Data Hackers

2021 - Entrevista Dados no Mercado Financeiro | Canal Programação Dinâmica

2020 - Palestra Da Física à Ciência de Dados | PUG Sergipe

2020 - Workshop sobre Ensino de Física com Python | PyLadies São Paulo

2020 - Palestra sobre Análise Numérica com Python | Python Brasil

Voluntariado

Líder de Comunidade | Comunidade Python Brasil | 12/2019 - Atualmente

Coordenadora | Python Nordeste 2022

No principal evento da comunidade Python no Nordeste, liderei a edição com foco em Diversidade & Inclusão, promovendo a participação de grupos minoritários tanto nas palestras quanto na organização. A conferência itinerante reuniu entusiastas de Python de toda a região, fortalecendo a comunidade e ampliando a representatividade no evento.

Cofundadora | PyLadies Sergipe

Fui uma das fundadoras do PyLadies Sergipe, uma comunidade que promove a participação e liderança de identidades femininas na comunidade open-source Python. Nosso foco é promover a Diversidade & Inclusão no setor de tecnologia no estado de Sergipe.