

Matheus Santos da Silva

Embaré, Santos – SP

(13) 99619-0428

msantos18.ms@gmail.com

<https://www.linkedin.com/in/mathssantos94/>

<https://www.matheussantos.com.br/>

OBJETIVO PROFISSIONAL

Cientista de Dados, Analista de Dados, Analista de BI

FORMAÇÃO ACADÊMICA

- Universidade Santa Cecília – Engenharia Civil (2014 – 2018)

EXPERIÊNCIA PROFISSIONAL

C-Freight Logistics

Período: 08/2024 – 10/2024

Cargo: Analista de Business Intelligence (Projeto Temporário)

Atribuições:

- Criação de dashboards e KPIs no Power BI, integrando múltiplas fontes de dados.
- Responsável por automação de relatórios e apresentação de insights com foco em metas e desempenho.
- Desenvolvimento de design no Figma;

Eleven Dragons

Período: 08/2022 – 03/2023

Cargo: UX/UI Designer

Atribuições:

- Desenvolvimento de uma plataforma web de streaming, liderando a construção do design system.
- Participação em reuniões ágeis com times de produto, design e tecnologia, facilitando o alinhamento de requisitos e entregas.
- Uso de Figma e FigJam;
- Responsável pela tradução de regras de negócio em interfaces intuitivas, comunicando-se com stakeholders e desenvolvedores de forma clara e objetiva.

PROJETOS PESSOAIS

Modelo de Avaliação de Risco de Crédito

(GitHub: github.com/Mathsnts18/CreditRiskModel | Aplicação: [StreamLit](#))

Descrição: Desenvolvimento de um modelo preditivo para identificar inadimplência em clientes de cartão de crédito, utilizando dados históricos de 30.000 contas.

Metodologias:

- Pré-processamento de dados: tratamento de desbalanceamento (22% de inadimplentes), correção de inconsistências (ex: remoção de colunas irrelevantes).
- Feature engineering: criação de categorias educacionais.
- Modelagem: Comparação de algoritmos (XGBoost, Random Forest, Regressão Logística) via MLflow, com registro de métricas e versionamento.
- Interpretabilidade: utilização de análise SHAP values para entender o funcionamento do modelo

Resultados:

- Identificação das variáveis mais impactantes.
- Análise financeira simulada indicou economia potencial de milhões ao direcionar intervenções para clientes de alto risco (custo de R\$1.300 por aconselhamento, com 70% de eficácia).

Ferramentas: Python, Pandas, Scikit-learn, XGBoost, MLflow, StreamLit

Predição de Preços de Apartamentos em Santos

(GitHub: <https://github.com/Mathsnts18/SantosApartmentPricing> | Aplicação: [StreamLit](#))

Descrição: Projeto de Ciência de Dados End-to-End com dados reais do portal Zap Imóveis, com foco em prever o preço de apartamentos na cidade de Santos/SP.

Metodologias:

- Web scraping com Selenium para coleta automatizada dos anúncios de imóveis
- Análise exploratória dos dados com identificação dos principais fatores que influenciam o preço.
- Tratamento de dados: remoção de duplicados, tratamento de valores ausentes, conversão de tipos e Winsorização de outliers.
- Feature engineering: criação de variáveis binarização de categorias e

one-hot encoding.

- Pipeline de Machine Learning com etapas encadeadas (RFE, StandardScaler, OneHotEncoder, etc.).
- Avaliação de modelos como Regressão Linear, Random Forest, LightGBM e CatBoost com validação cruzada e monitoramento via MLflow.
- Deploy do modelo final (CatBoostRegressor) com interface interativa no Streamlit.

Resultados:

- O modelo final apresentou MAE médio de ~R\$ 160 mil na validação cruzada e R\$ 163 mil no teste.
- App interativo capaz de realizar inferências em tempo real com base nas características de um imóvel.

Ferramentas: Python, Pandas, Scikit-learn, CatBoost, MLflow, StreamLit, Selenium, Matplotlib, Seaborn

CURSOS EXTRACURRICULARES

- Formação Data Science - Preditiva
- Formação Data Analytics – Preditiva
- Avançando em Data Science com Python – Alura
- Estatística do Básico ao Avançado (E.B.A) – Renata Biaggi
- Formação Estatística com Python – Alura
- Power BI para Análise de Dados – Preditiva
- Excel para Análise de Dados – Preditiva

HABILIDADES

Inglês: Intermediário

Excel: Intermediário

Google Sheets: Intermediário

Pacote office: Intermediário

Power BI: Intermediário

SQL: Intermediário

Python: Intermediário

Estatística: Intermediário

Machine Learning: Intermediário

Git/Github: Intermediário