

Ana Beatriz Carvalho | Cientista de Dados Sênior com foco em Machine Learning e soluções analíticas

Av. Paulista, 1000, São Paulo, São Paulo, Brasil, 01310-100

☎ +55 11 98765-4321 • ✉ ana.carvalho@datainsight.com.br

🌐 <https://anabeatrizcarvalho.com.br>

🌐 @anabeatrizcarvalho • 🐙 @anabeatriz-data • 📧 @anabeatrizcarvalho

Informações Básicas

- Cientista de Dados com mais de 8 anos de experiência em modelagem preditiva, mineração de dados e inteligência de negócio.
- Especialista em desenvolver soluções de machine learning para os setores financeiro e de varejo, gerando redução de custos e aumento de receita.
- Sólida vivência em Python, SQL, computação em nuvem (AWS, GCP) e ferramentas de visualização de dados.
- Experiência na liderança de times ágeis, na definição de indicadores e na comunicação de resultados para públicos não técnicos.

Formação

Universidade de São Paulo (USP)

Mestrado, *Ciência de Dados*, Nota: 9.1

<https://www.usp.br/>

jan. de 2015–dez. de 2017

- Dissertação sobre seleção de atributos e interpretabilidade de modelos preditivos.
- Atuação como monitora nas disciplinas de estatística e aprendizado de máquina.

Cursos: Aprendizado de Máquina Avançado, Estatística Computacional, Mineração de Dados, Modelagem Preditiva, Visualização de Dados

Universidade Estadual de Campinas (Unicamp)

Bacharelado, *Estatística*, Nota: 8.7

<https://www.unicamp.br/>

fev. de 2010–dez. de 2014

- Iniciação científica em modelos de sobrevivência aplicados à saúde pública.
- Participação em projetos de consultoria estatística para empresas locais.

Cursos: Inferência Estatística, Modelos Lineares, Amostragem, Análise Multivariada

Experiência Profissional

DataWise Soluções em Dados

Cientista de Dados Sênior

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

mar. de 2021–Atual

- Lidero o desenvolvimento de modelos de previsão de churn e propensão de compra para clientes de varejo e finanças.
- Implementei um pipeline de MLOps que reduziu em 40% o tempo de deploy dos modelos em produção.
- Criei um framework de testes A/B e de validação offline que melhorou a confiabilidade dos experimentos.
- Supervisiono cientistas de dados júnior e contribuo com a definição da arquitetura de dados.

Palavras-chave: Machine Learning, MLOps, Python, AWS, Airflow, Testes A/B

VarejoTech

Cientista de Dados

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

jan. de 2018–fev. de 2021

- Construí modelos de precificação e segmentação de clientes que elevaram a margem operacional em 12%.
- Desenvolvi painéis analíticos em Power BI para acompanhamento de metas e detecção de anomalias.
- Automatizei a coleta, limpeza e integração de dados de vendas e estoque com Python e SQL.

Palavras-chave: Segmentação, Preços, Power BI, SQL, Python, Previsão de Demanda

Grupo Financeiro Beta

Analista de Dados

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

jan. de 2016–dez. de 2017

- Atuei na construção de relatórios gerenciais de risco de crédito e inadimplência.
- Realizei análises exploratórias para apoiar a área de crédito na definição de políticas de concessão.
- Elaborei indicadores para comitês executivos e apresentações de resultados.

Palavras-chave: Risco de Crédito, Relatórios Gerenciais, SQL, Excel Avançado, Análise Exploratória

Idiomas

Português: Nativo ou Bilíngue

Inglês: Proficiência Profissional Fluente

Espanhol: Proficiência Profissional Limitada

Palavras-chave: Português do Brasil

Palavras-chave: TOEFL iBT 110

Palavras-chave: Leitura técnica

Competências

Machine Learning: Especialista **Palavras-chave:** Python, Scikit-learn, TensorFlow, PyTorch, XGBoost, LightGBM, MLflow, Feature Store

Estatística: Especialista **Palavras-chave:** Inferência Bayesiana, Séries Temporais, Testes A/B, Modelos Lineares, Modelos de Sobrevida

Engenharia de Dados: Avançado

Palavras-chave: SQL, Apache Spark, Airflow, dbt, Docker, AWS, GCP

Visualização de Dados: Avançado

Palavras-chave: Matplotlib, Seaborn, Plotly, Power BI, Looker

Prêmios

Prêmio ABRADe de Inovação em Dados

Associação Brasileira de Análise de Dados

nov. de 2023

Reconhecimento pelo projeto de MLOps aplicado à prevenção de churn em uma base com mais de 5 milhões de clientes.

Destaque Acadêmico

Universidade Estadual de Campinas

dez. de 2014

Reconhecimento pela excelência no desempenho do curso de Estatística entre os concluintes daquele ano.

Certificados

AWS Certified Machine Learning - Specialty

Amazon Web Services

<https://aws.amazon.com/certification/>

jun. de 2022

TensorFlow Developer Certificate

Google

<https://www.tensorflow.org/certificate>

mar. de 2021

Publicações

Revista Brasileira de Inteligência Computacional

Previsão de inadimplência em carteiras de crédito usando aprendizado de máquina

<https://www.rbic.org.br>

set. de 2020

○ Apresenta uma comparação entre regressão logística, florestas aleatórias e redes neurais para predição de inadimplência.

○ Discute estratégias de balanceamento de classes e interpretabilidade de modelos em cenários reais de crédito.

Referências

Ricardo Nogueira

Gerente de Dados na DataWise Soluções em Dados

+55 11 99876-5432

ricardo.nogueira@datawise.com.br

Ana Beatriz combina sólida formação estatística com capacidade de liderança e comunicação. Foi fundamental para a evolução dos nossos modelos de machine learning.

Projetos

Plataforma de MLOps para Previsão de Churn

Pipeline end-to-end que combina automação de treinamento, validação e deploy de modelos de churn.

<https://github.com/anab>

- Utiliza Docker, Airflow e MLflow para orquestrar as etapas do ciclo de vida do modelo.
- Reduziu o tempo de atualização dos modelos de semanas para horas no ambiente de produção.

Palavras-chave: MLOps, Docker, Airflow, MLflow, Churn

Modelo de Séries Temporais para Previsão de Vendas

<https://github.com/anabeatr>

Projeto de estudo com modelos ARIMA, Prophet e redes LSTM aplicados a uma base pública de vendas.

- Análise comparativa de acurácia entre abordagens clássicas e neurais.
- Desenvolvimento de um dashboard interativo em Streamlit para exploração dos resultados.

Palavras-chave: Séries Temporais, ARIMA, Prophet, LSTM, Streamlit

Painel de Indicadores de Negócio

<https://github.com/anabeatr>

Dashboard criado com Power BI para monitoramento de metas comerciais e qualidade dos dados.

- Consolidou mais de 15 fontes de dados em um único painel gerencial.
- Facilitou a tomada de decisão semanal dos times comerciais.

Palavras-chave: Power BI, Indicadores, DAX, SQL

Interesses

Leitura: Ficção científica, Não ficção

Esportes: Corrida, Yoga

Voluntariado

Mulheres em Dados

<https://www.mulheresemdados.org.br>

Mentora voluntária

jan. de 2020–Atual

- Mentoria de carreira para mulheres que estão ingressando na área de dados.
- Criação de conteúdos sobre machine learning e estatística para a comunidade.

Ciência de Dados para o Bem

<https://www.datascienceforgood.org.br>

Voluntária em projetos sociais

mar. de 2019–dez. de 2022

- Atuação em projetos de análise de dados para organizações sem fins lucrativos nas áreas de educação e saúde.