

Bruno de Sousa Donato

+55 (19) 98339-6071 – bruno.sousadonato@gmail.com

<https://www.linkedin.com/in/bruno-de-sousa-donato/> – <https://github.com/Bruno-Donato>

EXPERIÊNCIA PROFISSIONAL

- **CIENTISTA DE DADOS (10/2022 - ATUAL)**
Trabalho com empreendedores locais com análise de dados e business intelligence (métricas/KPIs), implementando pipelines e modelos, de regressão, classificação (KNN, SVM, Decision Tree/Árvore de decisão, Random Forest, XGBoost), clusterização (K-means), processamento de linguagem natural (NLP) utilizando webapps e dashboards para deployment de projetos. Usando criatividade para adequar o serviço ao perfil de cada cliente e tornar insights acessíveis e compreensíveis para orientar tomadas de decisões. Utilizo principais bibliotecas de análise e visualização de dados (Pandas, Numpy, Matplotlib, Seaborn, Plotly), ML/IA (Scikit-Learn, XGBoost, Scipy, StatsModels, Optuna, Spacy, NLTK, Gensim, REGEX, Optuna) e deploy (Streamlit, MLFlow)
- **CONSULTORIAS ACADÊMICAS (01/2020 - ATUAL)**
Consultorias de trabalhos acadêmicos para elaboração e análise estatística exploratória/descriptiva e inferencial. Aplicação e conhecimento em testes de correlação, modelos lineares e não-lineares, teste chi-quadrado, testes paramétricos (teste-t, ANOVA) e seus correlatos não paramétricos. Aulas e tutorias de fundamentos de estatística e pesquisa para área da saúde.
- **FISIOTERAPEUTA/CONSULTOR (SUS - PAULÍNIA/SP) (06/2019 - 03/2022)**
Avaliação, diagnóstico e reabilitação de pacientes ortopédicos e dor crônica, e trabalho com equipes multi e interdisciplinares do sistema de saúde público. Criação, gerenciamento e análise de banco de dados do serviço de reabilitação da cidade de Paulínia e seus usuários, orientando políticas públicas de saúde, promovendo direcionamento apropriado para atendimento e cuidado precoce (redução de 50% do tempo de espera para avaliação/atendimento), através de processo de triagem e classificação, baseada em dados epidemiológicos e estatística local e implementação de atendimentos em grande escala (aumento de 20% de capacidade de atendimento).
- **PESQUISADOR (BOLSISTA) - UNICAMP (01/2017 - 08/2019)**
Atendimento e coleta de dados dos usuários do ambulatório de reabilitação do hospital de clínica, gerenciamento e análise (descriptiva e inferencial) de dados do setor e projetos associados, com 2 anos de experiência em otimização de fluxo de pacientes, coleta e armazenamento de dados para fins de pesquisa clínica/acadêmica.
- **FISIOTERAPEUTA - AUTÔNOMO (01/2014 - 12/2022)**
Avaliação, diagnóstico e reabilitação de pacientes ortopédicos e dor crônica.

EDUCAÇÃO

- **Data Science & Machine Learning – TERA (09/2022 - 07/2023)**
Aprofundamento banco de dados, análise e visualização de dados, computação paralela e distribuída (Databricks/Spark), pipelines, modelos supervisionados e não supervisionados, inteligência artificial (IA), deploy de projetos
- **Mestrado – Programa de Pós-Graduação FCM - UNICAMP (01/2017 - 08/2019)**
Mestre em Ciências, com aprimoramento em didática acadêmica, metodologia de pesquisa científica e estatística
- **Graduação em Fisioterapia – PUC-Campinas (01/2009 - 12/2013)**
- Iniciação Científica (Bolsista) - fundamentos de estatística e pesquisa clínica, assim como produções científicas
- Monitor de disciplina: Neuroanatomia, anatomia e cinesioterapia
- **Ensino médio – Colégio Osvaldo Cruz Campinas (2005)**

IDIOMAS

- Português: Nativo
- Inglês: Avançado - TOEFL (01/2020)
- Espanhol: Básico

HABILIDADES E INFORMAÇÕES ADICIONAIS

- Linguagem de programação: Python (Avançado), R (Intermediário), SQL/NOSQL(Intermediário)
- Ferramentas: GitHub (versionamento e repositórios), VS Code, Jupyter Notebook, Power BI (Intermediário)
- Web Scraping: Selenium, BeautifulSoup
- Microsoft Office: Excel (Avançado) e Word (Avançado)

CURSOS

- SQL for Data Science IBM - DB0201EN (Edx - 2022)
- Machine Learning with Python: A Practical Introduction IBM - ML0101EN (Edx - 2022)
- Data Science and Machine Learning Capstone Project IBM - DS0720EN (Edx - 2022)
- Visualizing Data with Python IBM - DV0101EN (Edx - 2022)
- Analyzing Data with Python IBM - DA0101EN (Edx - 2022)
- Predictive Modeling and Analytics - Regression (Greatlearning - 2022)
- Python for Data Science Project IBM - PY0220EN (Edx - 2022)
- Intro to statistics. Making decisions based on data (Udacity - 2022)
- Analyzing and Visualizing Data with Power BI (Edx - 2023)
- ML and MLOps 10x faster. Hands-on MLOps/MLFlow PyCaret (Udemy - 2023)