

Carta de Apresentação

Sou um Cientista de Dados com forte experiência em machine learning, otimização e MLOps, focado no desenvolvimento de soluções escaláveis de IA e na implantação de modelos em produção. Sou formado em Engenharia Eletrônica e mestre em Sistemas de Informação, com experiência acadêmica internacional no Japão, Reino Unido e Canadá. Essa trajetória diversificada me proporcionou habilidades analíticas avançadas, permitindo compreender e implementar rapidamente pesquisas de ponta, aplicando métodos científicos para resolver desafios complexos da indústria.

Na minha experiência profissional, trabalhei extensivamente com frameworks de machine learning e deep learning para treinar e ajustar modelos voltados para modelagem preditiva, detecção de anomalias e visão computacional. Minha expertise em MLOps me permite projetar e implementar pipelines completos, garantindo reprodutibilidade e escalabilidade. Utilizando ferramentas como MLflow, Docker e plataformas em nuvem como AWS e DigitalOcean, implantei modelos de machine learning com sucesso e os integrei a sistemas de produção via APIs.

Com uma base sólida em análise estatística e otimização, tenho um interesse especial em IA Generativa e LLMs, explorando ativamente seu potencial em aplicações reais. Estou motivado a contribuir para soluções inovadoras de IA e ansioso para colaborar com profissionais e organizações que impulsionam a inovação em ambientes de produção.

Experiência Profissional

- WEG**
Cientista de Dados

Jaraguá do Sul, Brasil
Jun. 2024 – Presente

 - **Manutenção Preditiva & Detecção de Anomalias:** Desenvolvimento e implementação de modelos de machine learning e deep learning para manutenção preditiva e monitoramento de condição de turbinas eólicas.
 - **MLOps Completo:** Desenvolvimento de um pipeline escalável e automatizado de MLOps utilizando MLflow, Docker e AWS, garantindo reprodutibilidade no treinamento, implantação e monitoramento contínuo de modelos em produção.
 - **Implantação de IA & Integração com Nuvem:** Integração de modelos de ML em produção via APIs, otimizando inferência de modelos para detecção de anomalias em tempo real e análise preditiva.
- Macnica DHW**
Cientista de Dados

Florianópolis, Brasil
Nov. 2022 – Out. 2023

- **Aprendizado Supervisionado:** Desenvolvimento de modelos de machine learning para classificação de dados e suporte ao desenvolvimento de sensores inteligentes.
- **APIs de Visão Computacional:** Criação e implantação de APIs para aplicações de visão computacional, incluindo uma API de anonimização que remove pessoas de vídeos.
- **Desenvolvimento de IoT:** Liderança no desenvolvimento de um produto IoT, incluindo a implementação do protocolo MQTT e armazenamento de séries temporais usando InfluxDB.

• **Universidade de Warwick, Warwick Business School** Coventry, Reino Unido
Pesquisa de Doutorado *Out. 2019 – Abr. 2022*

- **Área de Pesquisa:** Ranking & Seleção; Otimização Bayesiana; Aprendizado por Reforço.
- **Pesquisa:** Sob supervisão de Juergen Branke, trabalhei no aprendizado eficiente de políticas em problemas de decisão sequenciais, aprimorando a eficiência amostral do Monte Carlo Tree Search (MCTS).
- **Colaboração:** Trabalho conjunto com Prof. Chun-Hung Chen (George Mason University, EUA) em abordagens Bayesiana para tomada de decisão ótima.

• **Shinshu University** Nagano, Japão
Pesquisa de Mestrado *Abr. 2017 – Mar. 2019*

- **Área de Pesquisa:** Otimização Multiobjetivo; Algoritmos Genéticos; Computação Evolutiva.
- **Pesquisa:** Estudo do impacto das topologias do conjunto de Pareto no desempenho de algoritmos evolutivos multiobjetivo (MOEAs) e desenvolvimento de um novo método para melhorar sua eficiência.
- **Colaboração Internacional:** Trabalho com pesquisadores do Japão, México e França.
- **Reconhecimento:** Recebeu o prêmio **Student Best Paper Award** na conferência EMO-2019 (Michigan State University, EUA).

SSE Gridtech Curitiba, Brasil
• *Estágio* *Out. 2014 – Dez. 2015*
Engenheiro de P&D *Jan. 2016 – Ago. 2016*

- **Soluções para Smart Grid e IoT:** Desenvolvimento de um sistema de comunicação LoRaWAN de longo alcance para otimizar a infraestrutura de medição inteligente, reduzindo a dependência de GSM.

Educação

• **Shinshu University** Nagano, Japão
Mestre em Engenharia Eletrônica e Sistemas de Informação *Abr. 2017 – Mar. 2019*
◦ Bolsa **MEXT Monbukagakusho** (Governo Japonês)

• **Concordia University** Montreal, Canadá
Intercâmbio, Engenharia Elétrica *Ago. 2012 – Ago. 2013*
◦ Bolsa **Ciência Sem Fronteiras** (Governo Brasileiro)

• **Universidade Tecnológica Federal do Paraná** Curitiba, Brasil
Bacharelado em Engenharia Eletrônica *Jan. 2010 – Jul. 2016*

Expertise Técnica

Linguagens de Programação:	Python, C, C++, Bash (Unix Shell).
Frameworks de Machine Learning:	PyTorch, scikit-learn, XGBoost, FastAI.
MLOps & Implantação:	Docker, MLflow, Hydra, Git, CI/CD (GitLab, GitHub Actions), FastAPI.
Engenharia de Dados:	Dask, PostgreSQL, Redshift, InfluxDB, SQL.
Plataformas em Nuvem:	AWS (SageMaker, ECR, S3, EC2), DigitalOcean.
Otimização:	Otimização Bayesiana, Algoritmos Evolutivos Multiobjetivo.
Sistemas Operacionais:	Linux (Fedora, Ubuntu), Windows.

Idiomas

Português:	Nativo
Inglês:	Fluente
Japonês:	Conversação

Projetos

- **Pipeline de Machine Learning para Predição de Preços de Aluguéis de Curto Prazo**
 - <https://github.com/yurimarca/build-ml-pipeline-for-short-term-rental-prices>
 - Desenvolvimento de um pipeline de Machine Learning reproduzível para prever preços de aluguéis de curto prazo em Nova York, garantindo escalabilidade para atualizações semanais de dados.
 - Foram Implementados processos de ingestão, limpeza, validação e engenharia de atributos para preparar os dados para modelagem.
 - Treinamento e otimização de modelos de regressão Random Forest, aplicando ajuste de hiperparâmetros para melhorar a performance preditiva.
 - **MLflow** foi utilizado para rastreamento de experimentos e gerenciamento de modelos, além de integrar **Weights & Biases** para rastreamento e visualização de artefatos.
 - Projeto inclui um pipeline para re-treinamento contínuo com novos dados, facilitando a melhoria constante do modelo e sua implantação.

Publicações

- **Y. Marca**, H. Aguirre, S. Zapotecas, A. Liefooghe, B. Derbel, S. Verel, and K. Tanaka. Approximating Pareto set topology by cubic interpolation on bi-objective problems. 10th International Conference on Evolutionary Multi-Criterion Optimization (EMO 2019), Lecture Notes in Computer Science (LNCS), Michigan, USA. (**Best Student Paper Award**)

- **Y. Marca**, H. Aguirre, S. Zapotecas, A. Liefooghe, B. Derbel, S. Verel, and K. Tanaka. NSGA-II with Spline Interpolation on Bi-objective Problems with Difficult Pareto Set Topology. JPNSEC 2018 Symposium on Evolutionary Computation, Fukuoka, 2018. (**Young Research Award**)
- **Y. Marca**, H. Aguirre, S. Zapotecas, A. Liefooghe, B. Derbel, S. Verel, and K. Tanaka. MOEAs on Problems with Difficult Pareto Set Topologies. IEICE Shin-etsu Branch IEEE Session, Niigata University, 2018, p. 169. (**Young Research Award**)
- **Y. Marca**, H. Aguirre, S. Zapotecas, A. Liefooghe, B. Derbel, S. Verel, and K. Tanaka. Pareto dominance-based MOEAs on problems with difficult pareto set topologies. In Proceedings of the Genetic and Evolutionary Computation Conference Companion (GECCO '18). ACM, New York, NY, USA, 189-190.
- C. E. A. L. Rocha, **Y. P. Marca**, F. K. Schneider. Support Platform for Decision-Making in Research and Technological Development in Public Health. ESPACIOS (CARACAS), v. 39, p. 14-26, 2018.
- **Y. P. Marca**, S. Scholze. Proposta de Substituição da Comunicação GSM em Smart Grids por Rádios de Longo Alcance. XXXIII Simpósio Brasileiro de Telecomunicações, 2015, Juiz de Fora, MG. Anais Completo da Programação Técnica, 2015.
- **Y. P. Marca**, C. E. A. L. Rocha, B. Schneider Jr , F. K. Schneider. Plataforma de Apoio ao Processo Decisório em Pesquisa e Desenvolvimento Tecnológico em Saúde. Congresso Brasileiro de Engenharia Biomédica, 2012, Porto de Galinhas. ANAIS - CBEB 2012, 2012.
- M. P. Krause, D. M. Nakato, **Y. P. Marca**, F. K. Schneider. Gerenciamento do Controle da Glicemia Utilizando um Aplicativo para Celular. Congresso Brasileiro de Engenharia Biomédica, 2012, Porto de Galinhas. ANAIS - CBEB 2012, 2012.