

Gabriel Przytocki

Cientista de Dados | Engenheiro de Machine Learning

E-mail: gabrielprzytocki@gmail.com

Github: <https://github.com/gprzy>

FORMAÇÃO ACADÊMICA

- **Especialização em Data Science e Big Data** Curitiba, Brasil
Universidade Federal do Paraná (UFPR) 2023
- **Bacharel em Ciência da Computação** Curitiba, Brasil
Pontifícia Universidade Católica do Paraná (PUCPR) Fev. 2019 - Dez. 2022

EXPERIÊNCIA

- **Junto Seguros** Curitiba, Brasil
→ **Cientista de Dados** Jan. 2022 - o momento
Design de produtos orientados a dados; Criação de modelos de Machine Learning;
Planejamento e execução de projetos; Identificação de oportunidades e otimizações;
Implementação de Sistemas de Recomendação; Modelagem de *churn* (rotatividade de clientes);
Implementação de modelos em produção; Desenho de arquitetura dos modelos de ML;
Melhores práticas de MLOPs; Comunicação com *stakeholders*;
- **Engenheiro de Dados Pleno** Out. 2021 - Fev. 2022
Criação de processos de ETL; *Pipelines* de processamento de dados;
Modelagem de dados com SQL Server e T-SQL; Agendamento e automações com Airflow;
ETLs de dados não estruturados (PDFs); *Web scraping* de dados com selenium e bs4;
Extração de textos em PDFs; Aplicação de OCR com pytesseract;
- **Pontifícia Universidade Católica do Paraná** Curitiba, Brasil
→ **PIBIC Undergraduate Researcher** Jul. 2021 - Jul. 2022
Modelagem de Séries Temporais Financeiras; Indicadores de Análise Técnica;
Previsão de ações com redes CNN e LSTM; Previsão de ações com modelos estatísticos;
Modelos ARMA e ARIMA; Comparação de algoritmos (Teste Wilcoxon);
- **PIBIC Undergraduate Researcher** Ago. 2020 - Ago. 2021
Modelagem de Séries Temporais Financeiras; Comparação de algoritmos (Teste Wilcoxon);
Previsão de ações com redes CNN e LSTM; Previsão de ações com modelos estatísticos;
- **Mirum Agency** Curitiba, Brasil
→ **Engenheiro de Dados** Nov. 2020 - Out. 2021
Fluxo de processamento de dados (ETLs); Cloud Functions, Cloud Storage e BigQuery;
Automatização de processos com Python; Integração de APIs RESTful;
Implementações com Google Cloud Platform; Arquiteturas *Datawarehouse* e *Data Lake*;
- **Software Developer Intern** Mar. 2020 - Nov. 2020
Desenvolvimento *front-end*; Implementações de *e-mail marketing*;
Javascript, HTML5 e CSS3; *Frameworks*: Bootstrap e JQuery;
- **Viasoft Korp** Curitiba, Brasil
→ **Software Developer Intern** Jul. 2019 - Dez. 2019
Desenvolvimento *back-end* em Delphi; *Patterns* MVVM e MVC;
Integrações com banco de dados (SQL Server); Desenvolvimento com testes unitários;

IDIOMAS

Português: Fluente ou Nativo

Inglês: Nível Avançado

Espanhol: Nível Básico

PROJETOS

- **Avaliação do Desempenho de Técnicas de Deep Learning para Classificação do Comportamento Temporal de Ações:** Relatório concluído de iniciação científica, referente a implementação e avaliação de técnicas de redes neurais (CNN e LSTM) em comparação com o modelo estatístico Autorregressivo (AR) na previsão em Séries Temporais Financeiras. → [Link para GitHub](#)
- **Previsão do Valor Futuro de Ações a Partir do Uso de Indicadores Técnicos:** Relatório concluído de iniciação científica, referente a implementação de combinações otimizadas por algoritmos genéticos de Indicadores Técnicos na previsão de ações em Séries Temporais Financeiras, juntamente com redes CNN e LSTM. → [Link para GitHub](#)
- **Aplicação de Fatores de Decaimento em Sistemas Adaptativos de Recomendação:** Trabalho de Conclusão de Curso (TCC) em andamento, referente ao curso de Bacharelado em Ciência da Computação. Aplicações de métodos de esquecimento (*forgetting*) de informações obsoletas nos algoritmos adaptativos (*stream*) ISGD e IBPRMF.

COMPETÊNCIAS E TECNOLOGIAS

Linguagens: Python, R, Julia, C, C++, Rust	Versionamento: git, GitHub
Machine Learning: scikit-learn, tensorflow	Estatística: scipy, numpy, pandas
MLOPs: AWS SageMaker, mlflow	Testes Unitários: pytest, unittest
Manipulação de Dados: numpy, pandas, polars	Big Data: pyspark, Spark, HDFS
APIs REST: flask, django, fastapi, streamlit	Cloud: AWS, GCP
Conteneirização: docker, kubernetes	Visão Computacional: cv2, pytesseract
Sistemas Operacionais: Linux, Windows	Séries Temporais: statsmodels, pycaret
Banco de Dados: SQL, BigQuery, NoSQL	SGBD: SQL Server, MySQL

FORMAÇÃO COMPLEMENTAR

- *Machine Learning Specialization* - (Stanford, Coursera);
- *Deep Learning Specialization* - (DeepLearning.AI, Coursera);
- *Analyst Developer* (Certificado Profissional - Pontifícia Universidade Católica do Paraná);
- *Data Scientist* (Certificado Profissional - Pontifícia Universidade Católica do Paraná);
- *Architecting with Google Compute Engine* (Coursera);
- *Elastic Google Cloud Infrastructure: Scaling and Automation* (Coursera);
- *Essential Google Cloud Infrastructure: Core Services* (Coursera);
- *Essential Google Cloud Infrastructure: Foundation* (Coursera);
- *Google Cloud Platform Fundamentals: Core Infrastructure* (Coursera);
- *Reliable Google Cloud Infrastructure: Design and Process* (Coursera);